PREDICTIVE PROCESSING DURING NOVEL WORD LEARNING: ERP MEASURES OF VOWEL HARMONY

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

NEUROSCIENCE

By

Berrak Müftüoğlu

June 2023

Predictive processing during novel word learning: ERP measures of vowel harmony
By Berrak Müftüoğlu
June 2023

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Hacı Hulusi Kafalıgönül(Advisor)

Miri Besken

Bilal Kırkıcı

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

PREDICTIVE PROCESSING DURING NOVEL WORD LEARNING: ERP MEASURES OF VOWEL HARMONY

Berrak Müftüoğlu
M.S. in Neuroscience
Advisor: Hacı Hulusi Kafalıgönül
Co-Advisor:
June 2023

This project investigated the effect of Turkish vowel harmony (TVH) on the acquisition of multi-syllabic words via cross-situational word learning (CSWL). CSWL is an implicit learning paradigm requiring learners to statistically track word-referent pairs across ambiguous conditions. Although there are many studies have examined certain lexical conditions that span word boundaries, within-lexical unit characteristics, such as TVH, have not been examined. We conducted 2 experiments; Experiment 1 is an internet-based experiment with Native Turkish speakers and speakers of other languages, and Experiment 2 is an EEG experiment focusing on N400 and LPC ERPs with Native Turkish speakers. In experiment 1, participants received training and recognition test phases. During training, participants were presented with novel pseudowords, both harmonious and disharmonious for TVH, with multiple referent pictures. In the recognition test, participants heard a pseudoword and were asked to select the correct referent for it. The results of the first experiment did not show a significant difference. In experiment 2, learning was additionally measured by a semantic judgment task (SJT). In the SJT, participants were presented with either a learned pseudoword or a real Turkish word, then with another real Turkish word. Their task was to indicate whether the two presented words were from the same semantic category. The recognition test results showed high accuracy, implying successful learning, with no differences among TVH conditions. The SJT results were modulated by both the novelty of the words (pseudowords vs. real words) and the harmony conditions in both behavioral and ERP measures. Specifically, harmonious words lead to N400 responses similar to real words, with no such effect on the disharmonious words. These results suggest that the phonological rules of participants' native language had an impact on semantic consolidation.

Keywords: Cross-situational word learning, Vowel harmony, Backness harmony, Flatness Harmony, N400, LPC, ERP.

ÖZET

YENİ KELİME ÖĞRENİMİ SIRASINDA TAHMİNE DAYALI İŞLEME: ÜNLÜ UYUMU OLAYA İLİŞKİN POTANSİYELLERİ

Berrak Müftüoğlu
Sinir bilim, Yüksek Lisans
Tez Danışmanı: Hacı Hulusi Kafalıgönül
İkinci Tez Danışmanı:
Haziran 2023

Bu proje Türkçe ünlü uyumunun (TÜU) durumlar arası kelime öğrenimi (DAKÖ) ile çok heceli kelime öğrenimi üzerindeki etkisini incelemektedir. DAKÖ belirsiz koşullarda kelime - referans çiftlerini istatistiksel olarak izlemelerini gerektiren bir örtük öğrenme paradigmasıdır. Literatürde DAKÖ ile kelimeler arası farkları inceleyen birçok çalışma olmasına rağmen, TÜU gibi sözcük içi birim özellikleri incelenmemiştir.

Bu çalışma kapsamında 2 deney yapılmıştır. 1. Deney internet üzerinden yapılmış olup ana dili Türkçe olan bireyler ve ana dilinde ünlü uyumu bulunmayan bireyler arasındaki farkı incelemiştir. 2. Deney bir elektroensefalogram (EEG) deneyi olup ana dili Türkçe olan katılımcılarla N400 ve LPC olaya ilişkin potansiyellerini (OİP) incelemiştir. 1. Deneyde alıştırma ve tanıma testi aşamaları uygulanmıştır. Alıştırma aşamasında katılımcılara Türkçe ünlü uyumuna hem uygun hem uygun olmayan yeni sözde kelimeler ve referans fotoğrafları ikili olarak sunulmuştur. Tanıma testi sırasında, katılımcılardan yeni sözde kelime duyduktan sonra o kelimeye ait olan referans fotoğrafını seçmeleri istenmiştir. 1. Deneyde ne ünlü uyumuna göre ne de katılımcıların ana diline göre anlamlı bir sonuç bulunamamıştır. 2. Deneyde öğrenme, 1. Deneye ek olarak anlamsal yargılama aşaması ile ölçülmüştür. Anlamsal yargılama aşamasında katılımcılar alıştırma aşamasında duydukları kelimelerden birini ya da Türkçe bir kelime duymuşlardır ve sonrasında başka bir Türkçe kelime okumuşlardır. Katılımcılardan duydukları ve okudukları kelimenin aynı anlamsal kategoriden olup olmadıkları belirtmeleri istenmiştir. 2. Deneyde katılımcılar tanıma testinde TÜU koşulları arasında hiçbir fark olmaksızın başarılı bir öğrenme göstermişlerdir. Anlamsal yargılama

aşaması sonuçları kelimelerin yeniliği (sözde kelimelere karşı gerçek kelimeler) ve TÜU koşullarına göre hem davranışsal hem de OİP ölçümlerinde farklılık göstermiştir. Spesifik olarak ünlü uyumuna uyan kelimeler farklı anlamsal kategorilerden olduğu zaman gerçek kelimelere benzer bir şekilde daha büyük genlikli N400 sinyali oluşturmuştur. Aynı kategoriden ve farklı kategoriden olan kelimeler arasındaki fark ünlü uyumuna uymayan kelimelerde görülmemiştir. Bu sonuçlar, katılımcıların ana dillerinin fonolojik kurallarının anlamsal pekiştirme üzerinde bir etkisi olduğunu göstermektedir.

Anahtar sözcükler: Durumlar arası kelime oğrenimi, Unlu uyumu, Buyuk unlu uyumu, Kuçuk unlu uyumu, N400, LPC, OİP .

Acknowledgement

Firstly, I would like to thank my advisor Dr. Alba Tuninetti for her endless support and patient mentoring during my graduate studies. This thesis would not have come to an end without her encouragement, and I feel very fortunate to work with her as my thesis advisor.

I would like to thank Dr. Hulusi Kafalıgönül for accepting me as his student. He eased a very stressful time in my graduate degree with his support. I would like to thank Dr. Miri Besken and Prof. Dr. Bilal Kırkıcı for their time to be members of the examining committee and contributions to my thesis.

I appreciate the help of my research assistants, Ecesu Ülker, Ege Gür, İpek Gümüş, Ayaz Karadağ, Ekin Velioğlu and Berk Yüce during data collecting. It wouldn't be possible to find that many participants so fast without their help and data collection would be less fun without their friendship.

I am sincerely grateful to my friend Şeyda Evsen. During these two years, she was there for me in every problem, we laughed and cried together, and her help was invaluable. Her friendship made grad school much more enjoyable.

I am thankful for the unwavering support of Alkım Özduran, who let me talk about my thesis for hours without complaining. I am thankful for her consistent help, from voicing my stimuli to trying to find the correct hardware to use.

And lastly, I owe a special thanks to Banu Karadeniz for letting me undergo my first EEG training with her participation. Her trust in me in this procedure gave me the confidence to do the study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 L2 Learning

Language learning is difficult whether infants are learning their first language (L1) or adults are trying to learn a second language (L2). During infant language acquisition, the difficulty of L1 learning is the novelty of the input, as they are trying to learn the language without having any prior knowledge of the world. On the other hand, L2 learners, especially adult learners, might have more trouble with the production and pronunciation of the language despite their prior world knowledge because their dominant L1 knowledge interacts with the incoming L2 information. Although L1 and L2 learning are two different learning situations, there are some common elements to both. Both L1 and L2 learning require an understanding of fundamental elements of languages such as grammar rules, new words, and the integration of those elements into meaningful sequences, such as sentences. For L1 learning, the learning environment provides cues to establish the understanding of language elements and help to overcome the challenges in learning. However, the environment usually does not provide cues at a comparable level for L2 learning, which makes the learning process more challenging for L2 learners. The challenge for L2 learners is not only the lack of environmental cues but also inhibiting their automatic L1 responses (Darcy et al., 2016).

For L2 learners, the automatic activation of L1 characteristics during learning fundamentally affects the L2 learning process. The effect of L1 on L2 learning is noticeable in the perception and discrimination of the sounds. This is critical for L2 perception, and as such, theories on L2 sound perception discuss this factor extensively. One of the most known theories in focusing on L2 sound perception is the Perception Assimilation Model L2 (PAM-L2) (Best and Tyler, 2007). The PAM-L2 focuses on articulatory gestures for L1 sound and how L2 sounds are categorized within L1 distinct sound categories due to their similarities. According to PAM-L2, the problem occurs when two L2 sounds are not different enough to be categorized in separate L1 categories; if there is no distinct categorization for an L2 sound in the L1, learners have trouble differentiating those L2 sounds. This discrimination issue may be caused from not being able to categorize an L2 sound in a single L1 sound category. In such a case, a single L2 sound would not have an exact L1 sound match, but it would be similar to multiple L1 sounds. Another problem appears when two L2 sounds are categorized under one L1 sound category. This situation would make it difficult to perceive the difference between those two L2 sounds since they are both similar to one single L1 sound. In both situations, the similarity between the L2 and L1 sounds is an important characteristic for the discrimination of L2 sounds. If L2 sounds are similar to L1 sounds, it is more difficult to separate the sounds into distinct L1 categories.

Another model that focuses on similar L1-L2 differences in L2 perception and production is the Speech Learning Model (SLM) and its recently revised version (SLM-r) (Flege, 1995; Flege and Bohn, 2021). Compared to the original SLM, the revised version puts more emphasis on the experiences of an individual and rejects the idea of an age limit in the perception and production of the L2. Instead, it says that L2 learners will never have the mastery of a native speaker because their own L1 phonetic subsystems interact with the newly learned phonemes. According to SLM-r, the L1 and L2 sounds share a common phonetic space which leads to a high degree of cross language interaction. Therefore categorization of newly learned L2 phonemes depends on many factors such as cross-language phonetic differences between the L2 and L1 phonemes. For a new L2 phoneme category to arise, there should be a strong distinction between specific L2 phonemes and existing L1 phonemes. If L2 phonemes are similar to L1 phonemes, it is more difficult for the learner to create a specific phoneme category. One important difference in the SLM-r compared to the SLM is the importance of age during L2 acquisition. Although the previous version focused on early and late bilinguals differently, the SLM-r proposes that the learning ability is the same in both situations yet the quality and quantity of L2 input creates the difference.

In both models, the influence of L1 plays a crucial role in L2 perception. This effect is also supported by many studies showing that the L1 influence cannot be differentiated from L2 perception. One of the factors that would contribute to L2 phoneme discrimination and production is learners' mastery of L1 sounds, especially sound variability in the L1 (Kartushina and Frauenfelder, 2014). Research done with Spanish native speakers learning French as their L2 showed that, if the speaker has less variable production for their L1 sounds, they are more accurate in producing both categorized and uncategorized L2 sounds (Kartushina and Frauenfelder, 2014). If a person has less variability producing a L1 sound, they are more accurate in producing the L2 sound independent of whether that sound can be categorized under an L1 sound or not. Another factor showing the importance of L1 might be the L2 training. In a study examining the effect of L2 education, freshman and senior-year Italian students in English majors and people without formal English education completed oddity discrimination tasks, where participants were presented with three sounds, two of which are the same and the one is different and asked to indicate the different one (Grimaldi et al., 2014). The results did not show a significant difference between the participant groups which implies that even an excessive period of training and interaction with the language in the classroom environment does not improve L2 sound perception significantly. The result of this study emphasizes the influence of L1 characteristics on L2 and how permanent this influence is.

1.2 L2 Word Learning

Even without the interaction of L1 characteristics, learning a language is a difficult task. The phonetic challenges described above can be seen during word learning, a crucial first step in L2 learning overall, both in infants and adults. Word learning can be especially challenging not just because it requires phonetic knowledge but also because there's an additional conceptual act of associating words with concepts in the environment. People can still communicate without knowing all the formal grammar rules, although their sentences might sound funny; yet, without knowing the proper words, communication would be highly disrupted. L1 or native speakers are provided intense environmental cues during infancy which includes and prioritizes word learning (e.g., parents pointing at objects and naming them for their children). For L2 learning, the help of the environment is not always present since L2 learning is usually carried out in isolated classroom environments, and learners keep using their L1 outside of that limited learning period. On the other hand, L2 learners might benefit from their prior knowledge about the world as that knowledge would help to create connections. The learner's prior knowledge is important for both being able to integrate the available cues and create a context with semantic information. According to Gillette et al., 1999, having the contextual information of a word, knowing the structure of the sentences, and corresponding imaginable concepts enhance unknown word identification. If one of those aspects is not present in the environment, word identification is interrupted due to little knowledge about the concept (Gillette et al., 1999). It is not possible to provide all those factors in learning the L1 since the infant learner would not know how to interpret the structure; therefore, L1 learning requires a high number of environmental cues. However, for L2 learning, learners' existing L1 knowledge could help to create the optimal environment to provide those factors. Moreover, adding semantic knowledge from L1 would contribute to creating connections between the novel words and information by creating a context for learning (Angwin et al., 2014). When associated with semantic information rather than other known names, novel words are recognized better (Angwin et al., 2014). Associating novel words with semantic information boosts the learning process by generating a more engaging environment for learning.

As L2 word learning (i.e., mapping an existing concept to a new word) is an essential but difficult part of language learning, there are many theories about how word-referent mapping occurs. Word learning can occur with a fast mapping technique where every word is presented with a single referent, and the one-to-one connection is clear. Although this is an effective technique to learn new words, a one-to-one mapping is usually not possible in real life environments where the objects are not in isolation, especially during L1 learning. Therefore, it may not be an ecologically valid approach. Another alternative is the idea that word-referent mapping can occur across multiple events in ambiguous conditions; for example, hearing the word "dog" in both a dog park and in a farm, without explicit information. One way to experimentally test this idea is called cross-situational word learning (CSWL) (Yu and Smith, 2007; Smith and Yu, 2008). This learning paradigm does not require isolated word-referent pairings; instead, multiple occurrences of words with referents in ambiguous situations are present in a CSWL paradigm, making it more similar to naturalistic conditions. According to existing CSWL research, people keep track of the multiple co-occurrences of words and their possible referents in the environment. A word-referent match is established as the co-occurrence of the referent with that word increases. Successful learning via CSWL has been observed for both adults and infants in experimental conditions (Yu and Smith, 2007; Smith and Yu, 2008). For example, infants aged 12-14 months showed higher looking times to the correct target compared to the distractor after a short two words-two object association training (Smith and Yu, 2008). In an experiment with adults, university students were trained to associate one English pseudoword with the correct referent among two to four possible referent object pictures (Yu and Smith, 2007). In all conditions, participants chose the correct referent above chance level in a multiple-choice test. These results imply that learning words in ambiguous situations by tracking the co-occurrence of the referents is possible, successful, and has been argued to be more ecologically valid (i.e., akin to how people learn in the real world).

Although people are still able to learn the referents of novel words with CSWL in ambiguous situations, an increased level of ambiguity makes it more difficult to identify the correct referent (Mulak et al., 2019; Vouloumanos, 2008; Yu and Smith, 2007; Yurovsky and Frank, 2015). The ambiguity can occur by increasing the co-occurrence of the word with a specific distractor (Vouloumanos, 2008). For example, participants' accuracy decreased significantly if a distractor was presented in half of the trials of a word in the training, compared to when the distractor was presented only once in every six trials (Vouloumanos, 2008). Another situation that creates high levels of ambiguity would be the number of possible referents in trials. The increased number of possible referents in the environment, whether there are two or four objects when the word is heard, can increase ambiguity. People are more successful at identifying the correct mapping when there is a smaller number of referents in the environment, and as the number of referents increases, people start to show lower performance (Mulak et al., 2019; Yu and Smith, 2007; Yurovsky and Frank, 2015). In their experiment with university students, Yu and Smith (2007) showed that the recalling accuracy decreased from two possible referents to four possible referents with English pseudowords. Replicating those results, (Mulak et al., 2019) found the same learning pattern for two to four referents and extended the results by showing that unambiguous situations (i.e., there is only one referent per word) lead to the highest accuracy level. Situations that include a higher number of possible referents create a highly ambiguous environment for learning as learners have to keep track of more objects to find correct word-referent pairs. Considering ambiguity from these two aspects, if a word is presented multiple times with the same objects (e.g. the word "Grink" was presented with both a picture of a bear and a tree multiple times), the learning would be highly disrupted. This effect is more visible if the participants matched the words with one of the distractor objects instead of the real referents in the first encounters with the word (Yurovsky and Frank, 2015). When the referent that people chose in the first trials is not presented in the next ones, people's tendency to find the correct referent in the second trial decreases as the number of referents increases from two to eight (Yurovsky and Frank, 2015). Therefore, it can be said that there is a limit to learning with CSWL as high levels of ambiguity cause a significant reduction in learning, although learning is still possible.

In addition to the level of ambiguity presented experimentally, words' phonological similarity could be another challenge for novel word learning via CWSL. Research has shown support for novel word learning via CWSL that includes phonologically similar words, yet a clear performance deficit is seen when the words are highly similar (Escudero et al., 2016; Mulak et al., 2019; Tuninetti et al., 2020). In particular, minimal word pairs, which are separated by only one vowel or consonant, create confusion and affect learning. There is also a difference between vowel minimal pairs and consonant minimal pairs in learning. Learning vowel minimal pairs (deet-dit) is more demanding than consonant minimal pairs (bon-ton) and non-minimal pairs (bon-deet) (Escudero et al., 2016). In this experiment, Australian English speakers recalled vowel monosyllabic nonsense words with lower accuracy than consonant minimal pairs and non-minimal pairs (Escudero et al., 2016). In a later study, the difference in learning vowel and consonant minimal pairs via CSWL was replicated, and it was found that consonant minimal pairs were affected less by the increased ambiguity due to an increased number of referents compared to vowel minimal pairs (Mulak et al., 2019).

Additionally, L2 vowel categorization within L1 sounds could make learning certain minimal pair words easier or harder in the CSWL procedure. When L2 vowel pairs were categorized as difficult or easy according to how well they are distinguished in L1, novel words with easy pairs were learned with higher accuracy on the test, although both conditions showed above change performance (Tuninetti et al., 2020). In these experiments, Australian English monolingual speakers were tested on Dutch and Brazilian easy, difficult, and non-minimal pairs and showed above change learning performance for all conditions, yet difficult minimal pairs had the lowest performance (Tuninetti et al., 2020). Those results imply that second language learning is possible via CSWL paradigm, even for the words that learners L1 do not distinguish well.

1.3 Prediction within word learning and vowel harmony

So far, the research using the CSWL paradigm with minimal pair words shows evidence of word learning as learners perform above chance in all conditions even though the performance for minimal pairs was lower than non-minimal pairs. The difference between minimal and non-minimal pairs is not surprising as there is less competition between the lexical structures of the words in non-minimal conditions (Vitevitch et al., 1999). As the competition between the words increases, which happens when the neighborhood density is high as in minimal pairs (Vitevitch et al., 1999), it is harder to associate the target words and corresponding object correctly. Especially when the words to learn are one syllable words, which is the case in most CSWL research with minimal pairs, the only differentiation technique would depend on the competition of one lexical unit of the words. However, multiple syllable words would have a different structure than one syllable minimal pairs, and this change could alter word learning. As the linguistic structure of the words becomes more complex with the increase in the syllable number, some additional lexical constraints that were not present in one syllable words could arise. Those constraints would shape the words within a language and create language-specific rules. These language-specific rules establish the standard wording forms within a language, and the majority of the words in that language should obey those rules. From the learning perspective, these rules should add predictive characteristics to the language since there are clear rules for word-formation. For speakers of these languages with these types of rules, we would expect that they would have expectations about the structure of words and as such, would be able to predict or expect next lexical components. Therefore after encountering the first part of a word, they would expect certain components for it to be a word.

Because of the predictive characteristics of the lexical constraints, it can be expected that the listeners could respond to situations that do not obey those lexical constraints differently. Those words would be the deviants in the language since they are not showing the characteristics of the majority of the words. Speakers of that language might not expect the lexical structure of those words with their language knowledge. Research with standard words, defined as the words that follow lexical constraints, and deviant words, defined as words that do not follow the lexical constraints, shows that deviant words lead to a higher mismatch negativity (MMN) responses, which is an EEG signal elicited by unexpected stimuli (Scharinger et al., 2010). The experiment used vowel alterations in German, where the first vowel of some words changes in plural forms but in other words, the vowel stays the same. For example, in an experiment, participants heard the words stock-stöcke (stick-sticks) and stoff-stoffe (cloth-cloths), with and without the appropriate vowel alterations. While stöcke shows vowel alteration for plural form because the 'o' sound in the singular form changes to 'ö', stoffe does not; therefore, the front vowel is only predictive for the word stöcke. In the EEG results, higher amplitude MMN responses were seen when the word stoff was presented with vowel altered version stöffe, which is illegal in German. These results show that lexical constraints add an expectation for the next portion of the word, and there are neural mechanisms that index when the expectation is not met.

Another lexical constraint that likely has predictive aspects, but has been considerably less studied in psycho/neuro-linguistics is that of vowel harmony. Especially in languages with vowel harmony rules, such as Turkic and Uralic languages, the order of the vowels have a major impact on word formation. Although vowel harmony rules are language-specific and show variation among languages, most of the time, vowel harmony requires words to contain vowels from only one subcategory of vowels or to have a consistent vowel subcategory change according to a pattern. For example, if the language has front-back vowel harmony, such as Turkish or Finnish, the words in that language should contain either only front vowels or only back vowels; therefore, vowels with different articulatory-acoustic characteristics cannot be in the same word (see Table 1.1 for a summary of Turkish vowel harmony rules). However, vowel harmony rules might be more complex than just including vowels from one category. Flatness vowel harmony in Turkish is an example of a complex rule. In flatness vowel harmony, the vowel selection does not depend on only one characteristic of the vowels as happens in front-back vowel harmony but also includes height and roundness as constraining mechanisms. In the flatness vowel harmony, unrounded vowels (a, e, ı, i) can be followed by unrounded vowels only. However, if the first vowel is from the rounded category, the selection of the next vowel will depend on both round-unround and high-low characteristics of the vowels; as such, the next vowel will be either round and high (u, ü) or unround and low (a, e). Therefore, flatness vowel harmony shows a more complex pattern compared to front-back vowel harmony in Turkish. In either case, the majority of words follow both rules and speakers of vowel harmonic languages expect vowels in the, e.g., subsequent syllables of a multisyllabic word to conform to vowel harmony rules.

Table 1.1: Phonological feature specifications of Turkish vowel system

	Front		Back	
	Unround	Round	Unround	Round
High	i	ü	ı	u
Low	e	ö	a	o

Since vowel harmony rules give certain rules for word construction in vowel harmonic languages, similar to other lexical constraints (e.g., similar to Spanish words that cannot begin with an /s/ followed by a consonant), disharmonic words should cause responses similar to those mentioned above for deviant stimuli. If vowel harmony is perceived as similar to other linguistic rules in vowel harmonic languages, harmonic words should be the norm, or standard words, and words that do not follow vowel harmony patterns should be the exceptions, or deviant stimuli, that cause, for example, a higher-amplitude MMN response. Hence, we can assume that the MMN response will be seen in response to vowel harmony violations since speakers of vowel harmonic languages will have unmet predictions about the vowels. Some EEG research shows exactly this: higher-amplitude MMN responses for those lexical units that violate vowel harmony (Scharinger et al., 2011; Ylinen et al., 2016). In a study with Finnish vowel harmony, participants listened to Finnish pseudowords where the deviant words had either vowel harmony violations or were phonologically legal, including both vowel harmonic words and words with a neutral vowel, as expected in Finnish vowel harmony (Ylinen et al., 2016). EEG recordings showed that larger MMN responses were seen as a response to the violation of vowel harmony compared to the phonologically legal deviants (Ylinen et al., 2016). Similar MMN responses to vowel harmony violations were also observed with Turkish participants for Turkish vowel harmony (Scharinger et al., 2011). Different from Ylinen et al. (2016), Scharinger et al. (2011) presented vowels individually instead of within words. Results of both research studies show that native speakers are able to recognize illegal vowel harmony deviants during passive listening. These results strengthen the idea that vowel harmony adds an expectation for the next sound, resulting in listeners forming a prediction about the next vowel.

In addition to responding to the vowel harmony violations, speakers of a vowel harmony language might also recognize vowel harmony patterns in other situations. As speakers of those languages are sensitive to vowel co-occurrence patterns, they can also detect the vowel patterns and segment the strings according to vocalic dependencies (LaCross, 2015). In research with speakers of Khalkha Mongolian (a language with vowel harmony rules) and American English speakers, participants were presented with pseudoword strings. The pseudoword strings were made up of the repetition of twelve three-syllable 'words' which followed a vowel harmony pattern. After a training phase with word strings, participants were asked to indicate which word had been heard in training out of two possible choices; one of the options was a three-syllable "word," and the other option was another three-syllable "part-word", which were constructed by taking syllables from different words (e.g., the last syllable and first syllables of two words in the word string created the part-word); therefore, part words did not follow a vowel harmony rule (LaCross, 2015). According to Khalkha Mongolian vowel harmony rule, the vowel sequence in words was more probable, while the probability of a word having a vowel sequence of part words was low (LaCross, 2015). Khalkha Mongolian speakers were able to detect repeated vowel patterns which were shown by choosing the full words in the test, whereas English speakers did not show such a preference (LaCross, 2015). Khalkha Mongolian speakers' preference for words over part-words for both the harmonic (although harmony preference was seen only for one vowel subcategory) and disharmonic conditions shows that the characteristics of the native language are not only helpful for recognizing vowel harmony, but they are also prone to detect vowel patterns within a foreign language. Therefore, the effect of lexical constraints on speakers' perception is not limited to their own language with certain characteristics, but speakers can also adapt to new rules.

L2 word perception is also affected by L1 characteristics as vowel harmonic language speakers still depend on vowel patterns in word strings. Moreover, if a language has consistent vowel patterns, either harmonic or disharmonic, vowel harmonic language speakers can tell whether a novel word might belong to a learned language even when not explicitly informed about any vowel harmony rule (Altan, 2011). In a research study with Turkish speakers, participants were trained on an artificial language where two-syllable words showed vowel harmony or disharmony or only half of the words showed a harmony pattern (Altan, 2011). After a short training phase where participants read vowel harmonic or disharmonic pseudowords aloud, Turkish native speakers were presented with two words and asked to choose which one was similar to the language they had read (Altan, 2011). If the training phase only included words with harmony patterns (either harmonic or disharmonic), in the testing phase, participants categorized harmonic words with harmonic words and disharmonic words with disharmonic words (Altan, 2011). These results showed that participants not only recognized the rule for vowels, independently of whether the rule was the same or different from their native language, but they also generalized these rules to the artificial language. Therefore, the influence of an L1 vowel harmonic language is not limited to detecting the pattern within words; it also includes the application of those vowel harmony rules.

Although native speakers of vowel harmony languages show an advantage for noticing vowel constraints more easily, vowel harmonic wording could still add expectations (and as such, predictivity) for speakers of languages without vowel harmony rules. However, they do not have the same experience with vowel harmony that speakers of vowel harmonic languages have. The only way for speakers of languages without vowel harmony to gain this skill would be to have some experience with vowel harmony as a part of L2 learning. In that sense, the real question would be whether they can learn vowel harmony rules without an explicit explanation. There is some supporting evidence for speakers of non-harmonic languages learning vowel patterns when the pattern is a simple rule with one-vowel characteristics (e.g., front-back). In this situation, there was no harmony or disharmony differentiation, but participants were able to recognize the overall pattern in both harmony and disharmony. The experiment was done with French L1 participants and the vowel harmony was manipulated by adding an accent to the real French words either making the word harmonious or disharmonious for roundness harmony (Skoruppa and Peperkamp, 2011). After listening to a story either with harmonious or disharmonious accents in the training phase according to their conditions, participants were able to accurately choose the words that belonged to the harmonic/disharmonic accent that they listened to (Skoruppa and Peperkamp, 2011). However, this preference was only seen when the vowel harmony accent had a clear characteristic pattern, such as roundness. When the same procedure was used with more complex harmony rules, such as changing the rule according to certain characteristics of the vowel, similar to flatness harmony in Turkish, participants were no longer able to classify the correct words (Skoruppa and Peperkamp, 2011). Therefore, during language learning, it is easier to learn simple compared to complex vowel harmony rules, and simple rules can be learned without explicit teaching.

While the previous study focused on native language pattern detection, this same process might be different than in an L2. In research done on learning a new artificial language (similar to L2 learning), both adults and children showed a bias toward choosing vowel harmony after a short training with the new language (Do and Mooney, 2022; Finley, 2012). In the experiment with adults, participants were trained with pseudowords and suffixes where the suffix obeyed different vowel harmony rules depending on the conditions that they were assigned (Finley, 2012). Participants were exposed to stem-suffix pairs in the training phase; immediately after that, they received a multiple-choice test with two alternatives asking which of the alternatives might belong to the language they had listened to in the training phase. In all conditions, participants were able to choose the same vowel harmony pattern for new words above the chance level. The experiment with children was done with a slightly different design to understand whether children would be able to choose the dominant vowel pattern. Therefore, in the training phase, they were presented with words that had a dominant vowel harmony pattern; in one condition, the majority of the words were harmonious for a roundness harmony rule, and in the other condition, the majority of the words were disharmonious for the roundness harmony rule (Do and Mooney, 2022). A two-choice test was presented after the training, and the children were asked to choose which option explained an accompanying picture (Do and Mooney, 2022). Surprisingly, even the children who received training with the artificial language where most of the words did not follow the vowel harmony pattern, showed a bias towards choosing the words with vowel harmony in the test. This result supports that vowel harmony might be preferred by speakers of languages which do not show vowel harmony rules even when they are learning a non-harmonious language.

However, all the studies mentioned found a bias toward vowel harmony using one characteristic vowel harmony pattern: either height harmony or roundness harmony. A similar situation was also seen in Skoruppa and Peperkamp (2011) paper, where participants were able to recognize simple vowel harmony patterns but more complex patterns with multiple rules were difficult for them. Therefore, whether people can understand and generalize more complex or multiple rules remains to be seen. For example, in Finley's (2021) paper, participants' ability to learn and apply two different suffix types was compared; one suffix with vowel alterations according to height vowel harmony and the other one having one specific vowel in all cases without alterations. In the experiment, native English speaker participants were trained with a new language with two different types of suffixes. After the training, they received a two-choice test for four different item classes; training items that they saw, new items and suffixes with vowel alterations, new items and suffixes with non-changing vowels, and new items with both suffixes. They were asked to indicate which of the presented words belonged to the language they were presented with in training. Participants were good at choosing the correct options for old items and new items with the alternating suffix; however, especially in the test items with both suffixes, participants did not show better performance than the control group who did not receive any training. According to the results, it seems like when there is a vowel harmony pattern, participants can notice the pattern for changing suffixes; however, adding two types of suffixes together becomes complex.

Since these research studies imply that speakers of languages without vowel harmony can recognize vowel harmony to choose words after training, another question to ask might be whether they can adapt to vowel disharmony as a pattern too. To answer this question, a similar artificial language learning method was used in comparing harmony and disharmony patterns. In this experiment, native French speaker participants were trained with singular and plural words where plural French pseudowords were made with either harmonic or disharmonic suffixes, and then they were asked to choose the plural form of novel words. The participants who were trained with vowel harmony performed more accurately (Martin and Peperkamp, 2020). Similar results were also found with English participants; after a single pseudoword-suffix training, they were tested with double suffixes for plural and diminutive, testing either a harmony or disharmony rule (Martin and White, 2021). In both experiments, participants had higher accuracy in the vowel harmony condition. The studies comparing vowel harmony and disharmony provide evidence that accepting vowel disharmony as a rule is more difficult than vowel harmony. Such evidence would imply that learning vowel harmony could be easier than learning disharmony as a rule in the language. Moreover, since in these studies, vowel harmony was not explained to the participants, there might be a natural and general tendency to prefer vowel harmony over disharmony during naturalistic language learning.

1.4 EEG: N400 and LPC

On the other hand, all the studies mentioned measured participants' accuracy in choosing one or the other options. These kinds of behavioral responses, which require participants' overt responses, might not always give the full picture of the situation as the results depend on choosing an option. Processes like word learning require more changes in the brain such as semantic association building, consolidation, and building of memory traces which cannot be seen with behavioral measurement. In such cases, other techniques, such as event-related potential (ERP) measurements, which are time-locked measurements to certain stimuli, with EEG might result in a more accurate and deeper understanding of the participant's responses. The greatest advantage of ERP measurements is their high temporal resolution, which provides time-sensitive measurements of the activity. Therefore, by using ERPs, time-sensitive information about the changes in the neural signals in response to certain stimuli can be acquired. Some of the known language-related ERP signals that focus on semantic learning processes are N400 and late positive complex (LPC). These two ERP signals will be the primary focus of this project.

The N400 is a negative-going ERP response that usually peaks around 400ms after the onset of the stimuli. It is usually associated with lexical-semantic processing; a higher amplitude N400 ERP arises in response to incongruent and unexpected words. It was first found by Kutas and Hillyard (1980) in response to an improbable sentence ending. In that experiment, participants read several seven-word sentences. In the semantically inappropriate condition, a quarter of the sentences ended with a semantically incongruent word, and in the physically deviant condition, the last word was written with capital letters (Kutas and Hillyard, 1980). In the EEG measurements, a negative-going response peaked around 400ms in response to the semantic incongruity. This ERP pattern was not seen in physical deviants, such as the larger font size. Moreover, there was a correlation between the semantic incongruence and N400 ERP amplitude; strong mismatches (e.g., "He took a sip from the transmitter") led to a higher amplitude N400 response compared to moderate level mismatches (e.g., "He took a sip from the waterfall") (Kutas and Hillyard, 1980).

Although the N400 response was seen in the sentence condition in Kutas and Hillyard (1980) paper, sentence structure is not a necessity for the N400 response. Word pairs from the same or different semantic categories or violations within semantically related word sequences would also lead to an N400 response in the EEG. For example, in the third experiment of Nobre and McCarthy (1994), participants were presented with either semantically related (chalk-crayon) or unrelated (child-tree) prime-target pairs. In the ERP results, valid targets (i.e., semantically-related words following the first word of the pair), showed a lower amplitude N400 response compared to all other conditions. In this case, the more semantically related the second word was to the previous word was associated with a concurrent decrease in the N400 amplitude, showing that the semantic relatedness affects N400 amplitude regardless of sentence context. This characteristic of the N400 shows that it is not the place of the word that is important (e.g., the end of a sentence), but the context of where the word is used has the real impact (e.g., its semantic relation to other words around it). In other words, the expectation about a word in a specific context affects the N400 ERP amplitude. When the following word does not match participants' expectations, the amplitude of the N400 response increases, whereas when the word is congruent with expectations, a lower-amplitude N400 is seen.

As the N400 is connected to violations of semantic expectations, it can be assumed that presented words should have a semantic meaning for the participant. However, some research found an N400 signal in response to pseudowords, which are readable words without actual meanings (Nobre and McCarthy, 1994). Such activation shows that although the N400 reflects semantic context, if the presented stimuli conform to the phonotactics of real words in the tested language, there will still be a noticeable N400 response. On the other hand, this N400 activation does not appear to illegal or unreadable words (Nobre and McCarthy, 1994). To see the difference between different word types, including pseudowords and non-words, researchers conducted an experiment with a word list task. Participants were asked to indicate whether the presented word was from the given target semantic category (e.g. fruits and vegetables). In the results, no significant difference was found between the N400 responses for pseudowords and content words. Both pseudowords and content words had the highest N400 amplitude compared to function and nonwords. Unlike pseudowords, nonwords did not lead to an N400 and had significantly less negative than all other word types. Having such an effect for nonwords is expected: nonwords cannot constrain semantic information because they are illegal in the tested language.

This differentiation between non-words and pseudowords could also be helpful to understand second language word learning since most of the words in a second language are unknown and, except for cognates, do not have an immediate semantic link in the L1. Therefore, attributing semantic information to previously unknown L2 words would lead to changes in N400 responses. Research on L2 learning and the N400 response demonstrates that unknown L2 words show brain activity similar to nonwords and do not cause higher amplitude N400 components at the beginning of learning; however, the N400 component starts to arise as quickly as 63 hours after formal L2 instruction (McLaughlin et al., 2004). This shows that meaning acquisition while learning L2 can be very fast in changing the brain signal pattern after a couple of days.

A similar effect to L2 words can be seen with learning novel words either in a novel language or L1. Supporting that, N400 amplitude change after a short amount of training was also recorded with rare English words for native English speakers (Perfetti et al., 2005). The research was conducted with native English speakers and rare English words that the participants did not know. Consequently, those words have higher N400 amplitudes as they are legal words that do not fulfill the expectations of the participants (Perfetti et al., 2005). In the experiment, participants were given the training to learn the meaning of those rare words, and after the training trained rare words, untrained rare words and real words were matched with either correct or incorrect meanings. Trained rare words led to a higher N400 amplitude when they are presented with incorrect meaning, whereas such a difference between correct and incorrect meanings was not seen in untrained rare words. This shows that unless the learning is accomplished and the words are associated with meaning, the N400 response would not be sensitive to semantic mismatches. For that reason, N400 amplitude changes to congruent and incongruent relations are a way to measure word learning.

Although learning rare words gives an insight into the interaction of the N400 component and word learning, one factor to consider might be the effect of participants' language background. Even though rare words are much less frequent in the language, they could be very similar to more frequent words or participants might have known the rare words or encountered those words before. Such situations would make it difficult to diminish the effect of previous exposure to novel word learning. One way to decrease this effect could be using pseudowords instead of rare words as pseudowords would still obey the rules of the specific language but do not have associated meanings. In that aspect, pseudowords would allow more control over the stimuli as participants cannot be familiar with the words.

As it was mentioned, as long as the words are orthographically legal and readable, pseudowords can produce N400 ERP around the same amplitude as real congruent words. However, this effect was seen in experiments where participants were given a category name and asked to indicate whether the "word" presented on the screen was from that semantic category (Nobre and McCarthy, 1994). Therefore in that situation, the meaning of the presented word does not matter as long as it does not belong to that category. On the other hand, since pseudowords do not have meaning, the pseudoword-word pairs should not lead to a difference in N400 ERP amplitude at first. The ERP response should be similar to the second language or rare word learning such that, until the meaning is acquired, no difference in N400 response to mismatches would be seen. This change in N400 due to newly created associations has been recorded in word recognition studies in reading modality (Angwin et al., 2014; Bermúdez-Margaretto et al., 2018, 2019).

In Bermúdez-Margaretto et al. (2018, 2019), a pseduoword - meaning association was instantiated by presenting a picture of an object after half of the pseudowords while half of the pseudowords were followed by a hash mark. After some blocks of training, a decrease in N400 response for meaning associated pseudoword was observed, but such difference in N400 did not occur for pseudowords with hash marks. This effect is seen both in silent reading tasks where the participants passively read the words on the screen and in a task where they are asked to identify whether the presented word is a real word or not (Bermúdez-Margaretto et al., 2018, 2019). Importantly, the N400 response amplitude was the same for both picture-associated and non-associated pseudowords in the first block; yet, the N400 response for the picture-associated pseudowords was lower in the last block (Bermúdez-Margaretto et al., 2018, 2019). Considering the characteristics of the N400 component, this decrease in the N400 response could imply that an association between pseudowords and pictures was established and participants expected the corresponding picture to appear after hearing the word.

In another study with pseudoword learning, novel words were presented orthographically as names of aliens. The novel words were presented with an alien picture and either two semantic attributes or two other proper names (e.g. Zealt Timid Proud or Proach Mansey Smeath) (Angwin et al., 2014). In the testing phase, newly learned words (names of the aliens) were presented with either the associated alien picture in the congruent condition or another picture in the incongruent condition (Angwin et al., 2014). Results of the experiment showed higher amplitude N400 responses for incongruent pairs when the novel word was presented with semantic attributes, while that difference to congruent-incongruent words was not seen to novel words presented with other proper names (Angwin et al., 2014). Because semantic associations between words are the key to N400 amplitude modulations, a stronger association for the words with semantic attributes in addition to pictures leads to higher amplitude changes in the N400 as a response to incongruence. As illustrated by both studies, the difference in the N400 responses due to congruent-incongruent pairs starts to be seen after the association is established.

Although the N400 response for newly learned words starts to emerge after a relatively short period of exposure, some studies propose a difference between N400 responses for previously known familiar words and newly learned words. The general tendency for the newly learned words seems to be an increase in the anterior N400 activity compared to the familiar words (Mestres-Misse et al., 2007). In the experiment, participants were exposed to novel words in multiple sentences. After the training, related and unrelated word pairs for both novel and real words were presented. The EEG results showed N400 response for both novel words and real word pairs (Mestres-Misse et al., 2007). However, while the N400 effect for real words was more prominent on the right central/posterior region, novel words lead to an N400 effect on frontocentral regions.

Although there is a difference in anteriority of the N400 response for different stimuli types, the general activity pattern does not show a difference in lateralization for the different stimuli types. From the results of experiments looking at lateralization and N400 response, there does not seem to be a significant difference between the left and right hemispheres (Kandhadai and Federmeier, 2010). Contrarily, the LPC ERP response, the other ERP response we are focusing on in this paper which was also measured in the same experiment, showed an interaction between conditions and lateralization. According to Kandhadai and Federmeier (2010), there is a higher amplitude of the LPC in the right hemisphere with a higher semantic relation; however, this difference was not present for the right hemisphere.

The LPC is another ERP signal that can be seen after N400 ERP, around 500 to 700ms after the onset of stimuli, in word learning studies. The presence of an N400 is not required for the LPC component to arise. Unlike N400 ERP, the LPC is not a language-specific ERP component, but it is usually associated with explicit episodic memory traces (Rugg and Curran, 2007). According to Rugg and Curran (2007), an LPC-like signal is elicited in response to a recollection of a specific information which requires people to exactly remember items. For that reason, the LPC is known to play a role in encoding and retrieval of the novel word in word learning studies. From this aspect, the LPC component is usually associated with the activation of episodic memory traces which requires a more controlled process while the N400 is more automatic and appears when the learning is well-established with semantic information. For example, in a novel word learning study with rare real words, having a training session with the phonology of the words led to an increase in the LPC amplitude later in a silent reading task (Bermúdez-Margaretto et al., 2020). The change in the amplitude of the LPC would simply be a result of participants' familiarization with the word without any semantic activation as the words were not associated with any meanings. This difference on the LPC component in this experiment can be interpreted as the ease of activation of episodic memory traces after training for rare words. On the other hand, the N400 response was seen only after adding semantic components to the training phase. In the same paper, after receiving a training phase with both phonological and semantic information for rare words, a difference between trained words and other words wes only seen in the N400 component rather than the LPC in a silent reading task (Bermúdez-Margaretto et al., 2020). This difference in the EEG components was explained by saying that the additional semantic information enhances memory traces which end up leading to easier recruitment of the trained word, and in turn, this places less demand on episodic memory traces (Bermúdez-Margaretto et al., 2020). Supported by these results, the LPC component is seen to index a situation that requires an explicit recall of the meanings with access to episodic memory traces, whereas the N400 component is seen when a strong word association is formed, so it is easier for the person to recall that information.

Considering the relation between the LPC and N400 components, where one implies an easier recall process, it is possible to think there is an order in their appearance. LPC seems to appear in earlier periods of learning when recalling the meaning requires information from episodic memory whereas the N400 response is seen after the semantic information is connected to the word. Supporting this idea, studies focused on semantic consolidation have not found a robust N400 neither immediately after training (Bakker et al., 2015) nor after some offline consolidation period (Lei et al., 2022; Liu and Hell, 2020). On the other hand, the LPC response was found after a 24 hour consolidation period, although it was not present on the day of training (Bakker et al., 2015; Liu and Hell, 2020; Lei et al., 2022). In these studies, participants were given training for known words with their meanings and novel words which are matched with two-word meanings, such as 'blue cat' (Bakker et al., 2015; Liu and Hell, 2020). They were given training for two days: in each day, they were presented with different words, and after the second day of training, they were given a semantic judgment task where they should indicate whether the presented word which was trained is related to the real word presented after it (Bakker et al., 2015; Liu and Hell, 2020; Lei et al., 2022). This semantic judgment task was also repeated after an offline consolidation period of a week (Liu and Hell, 2020). In the results of all three papers, no N400 component difference was seen between congruent and incongruent novel words, whereas the LPC was present in response to incongruent semantic pairs for the first day's novel words in the first EEG task, and after the offline period, an LPC was seen as a response to incongruent pairs for both first- and second-day's novel words. The lack of an N400 component suggests that even after one week of consolidation, the recall of newly learned words did not become as automatic as existing word knowledge, and participants still relied on a more controlled recalling process. However, it is important to note that, an N400 difference was seen in novel words in Bakker et al. (2015), without involving congruency. The words which were learned on the first day elicited N400 responses similar to the real word on the second day, while the words presented on the second day were more negative.

1.5 CSWL and ERP components

As there is evidence of neural changes with explicit learning paradigms in previously mentioned studies, another way of looking at word learning by using ERPs would be investigating implicit learning paradigms such as CSWL. The CSWL paradigm has recently started to be researched with ERP methods (e.g., Angwin et al., 2022; Mangardich & Sabbagh, 2022). These studies have shown a robust N400 response for incongruent novel word-real word pairs with both adults (Angwin et al., 2022) and children (Mangardich and Sabbagh, 2022). In the study with adults, after a standard CSWL paradigm with training and recognition test, an additional semantic judgment task was run. In the semantic judgment task, participants were asked to indicate whether presented word pairs were from the same or different semantic categories. The results of the experiment showed a higher N400 negativity for the incongruent novel word pairs compared to congruent pairs. This pattern was similar to the real word -incongruent/congruent pairs that were presented as the control in the semantic judgment task (Angwin et al., 2022). The study with four year old children also used the CSWL paradigm but changed some tasks for children. Instead of a semantic judgment task, a Visual-Picture/Spoken-Word Matching Paradigm was used, where the trained words were presented with either the correct or wrong referent picture for that word. The results also showed a more negative N400 response for incongruent pairs (word - wrong referent picture). In both studies, such outcomes can be possible only when participants learned associations between the novel words - correct referents and were able to retrieve such information. This suggests that novel words can be learned and retained such that they elicit an N400 response when a semantic violation occurs.

1.6 Current Study

Considering the successful semantic acquisition via CSWL and an N400 presence in those studies, the current project broadens the topic by focusing on the effects of vowel harmony, a within-lexical constraints on CSWL. The aim of the current project is to investigate the effect of vowel harmony on word learning via CSWL by using ERP measurements. Specifically, the experiment focuses on Turkish vowel harmony aiming to see whether native Turkish participants would show a difference in learning harmonious and disharmonious words. Using ERP measurements, specifically the N400 and LPC components, enables investigation of the potential interaction between within-lexical unit phonological characteristics (e.g., vowel harmony) and semantic consolidation and retention mechanisms during language acquisition. The individual elements of this study, the CSWL paradigm, vowel harmony, and N400 and LPC components are well-examined concepts. However, the possible effect of vowel harmony on word learning via CSWL has not been examined, neither behaviorally or neurally.

To our knowledge, studies focusing on learning words with vowel harmony via ERP measurements is very limited, yet there is very valuable information. Research with native English speakers who learn Finnish showed that detection of vowel harmony in pseudowords does not happen directly after learning the rule, but it requires some experience with the vowel harmonic language (Osterhout et al., 2019). In the experiment, the researchers measured ERP responses to real Finnish words, legal vowel harmony pseudowords, and vowel harmony violated pseudowords in three different time points; the first month of learning, the fourth month of learning, and after the cessation of formal instruction (Osterhout et al., 2019). In the first month of learning, both legal pseudowords and vowel harmony violations led to an N400 response compared to real Finnish words. This shows that in the beginning of the learning, participants do not use vowel harmony as a way to differentiate words, because both legal and vowel harmony pseudowords were perceived similarly.

However, after four months of instruction, vowel harmony violations stopped leading to a N400 response, instead elicited a P600 response which is an ERP component that is associated with syntactic anomalies. That means L1 English speakers reached a point where they can detect Finnish vowel harmony within the words (Osterhout et al., 2019). As the N400 response arises for semantic anomalies and P600 for syntactic anomalies, the result of the experiment implies that learners recognize all words from their semantic units when they first start to learn. Therefore, participants were not using the vowel harmony rule to decide whether a word real, but they were only deciding according to the vocabulary they had. Moreover, learning those vowel harmony rules showed a gradual pattern correlating with the learning period. The more the learner has gained experience with Finnish, the better they become differentiating legal psedudowords from vowel harmony violations. Supporting that gradual learning for vowel harmony rules, the opposite pattern was seen for the regression period on the third measurement. The ones who had their last Finnish class more recently showed the P600 effect in response to vowel harmony violations, whereas as more time passed, the N400 effect started to be seen for both vowel harmonic and non-vowel harmonic pseudowords.

Although the previous paper explains the reaction to vowel harmony from the perspective of a native speakers of a non-harmonic language, the effect of vowel harmony on learning L2 for a speaker of vowel harmonic language is a different question. In the current study the focus was on the speakers of a vowel harmonic language, Turkish, and how having such a rule in the native language influence learning novel words. The previous literature showed that speakers of vowel harmonic languages have a tendency to vowel patterns in novel languages with behavioral measurements (Altan, 2011; LaCross, 2015). Moreover, research showed a difference for word categorization with simple vowel harmony rules and complex vowel harmony rules on French speakers, which is a non-harmonic language. Turkish gives a special case in that aspect by including both a simple vowel harmony rule, Backness Harmony, and a complex rule, Flatness harmony. Therefore, this study aims to expand the literature by adding how such a tendency to harmony patterns influences novel word learning considering both Backness and Flatness harmony rules. CSWL was chosen as the word learning paradigm rather than one-to-one paradigms like fast mapping to create a more naturalistic environment. Since, in the natural environment, people learn most of the words by encountering them multiple times in different conditions, an implicit paradigm like CSWL would create similar conditions.

The current project investigated three research questions through two experiments. The first research question is whether vowel harmony has an effect on word learning via the CSWL paradigm. We investigated this question in both experiments. In the first experiment, the focus was on both vowel harmony rules (Backness and Flatness) with native speakers of Turkish and speakers of non-harmonic languages. This internet-based experiment was prepared to create a base for the second experiment (ERP experiment). Considering the previous literature on CSWL, we expected participants to be successful in all words. We expected to see an effect of vowel harmony for the Turkish participants as the disharmonic pattern would be a violation for the rules of native language. On the other hand, we did not expect a differentiation according to vowel harmony rules for the speakers of non-harmonic languages as there are no such consistent patterns in their native languages. Word learning is operationalized with a recognition test in the internet-based experiment.

In the ERP experiment, a semantic judgement task (SJT) was used as a way to measure semantic consolidation in addition to the recognition test. In the SJT, participants were presented with novel pseudowords and real Turkish words auditorily followed by a orthographical real Turkish words. Participants were asked to indicate if the presented words (one auditory and one orthographical) were related by being from the same semantic category. All the participants in the ERP experiment were native Turkish speakers. Different than the internet-based experiment, the only between-subject factor was the Backness harmony, and the Flatness harmony was used as a within-subject factor. This change was done to control the characteristics of vowels in each between-subject condition.

We expected to see a difference between harmonic and disharmonic words in the SJT. The hypotheses for the effects of Backness harmony and Flatness harmony in the SJT were separate. As the Backness harmony was a between-subject factor, considering the previous literature where participants were successful at adapting the vowel harmony rule when the rule is consistent (Altan, 2011; LaCross, 2015), we did not expect to see a significant difference in Backness harmony on the performance. However, since the Flatness harmony was a within-subject factor and participants heard both harmonic and disharmonic words, we expected to see an effect of Flatness harmony on semantic consolidation shown in the accuracy. Considering Flatness Disharmonious words cause a violation for the participants' native language, we expected them to have a difficulty learning disharmonious words. This difficulty would be seen as a lower accuracy rate in the SJT.

The other two research questions were about the ERP results by asking whether newly learned words via CSWL would create semantic process related ERPs, N400 and LPC and whether vowel harmony would have an effect on these ERP signals. For the first question, we expected to see an N400 difference between related and unrelated word pairs for both Real and Novel words in the semantic judgment task implying participants successfully learned the words. These results would also be the replication of the results of Angwin et al. (2022) and Mangardich and Sabbagh (2022), where N400 ERP was present in response to unrelated semantic associations for newly learned words.

The effect of vowel harmony is a more interesting question for this study. Similar to the hypothesis for the behavioral results, we do not expect to see a difference between Backness harmonious and Backness Disharmonious conditions as it was a between subject factor. On the other hand, we have multiple hypotheses for the Flatness Harmony factor depending on the strength of semantic consolidation since the literature showed N400 or LPC increase in different situations. Taking the differences in the literature to account, we hypothesized that in case of strong semantic consolidation, the difference would be present in the N400 signal implying an more automatic processing. However, if the recalling is less automatized, the difference between harmonious and disharmonious words would appear in the LPC signal. I will explain the hypothesis for each possibility in more detail below.

According to literature, incongruent word - referent pairs with newly learned words first generated an LPC signal and the signal turns to an N400 signal as the recalling becomes more automatic. Therefore, in the strong semantic consolidation process with an automatic recalling, we expect to see a difference for harmonious and disharmonious words in N400 time window depending on the relatedness. We expected to see the largest N400 difference between Real Word - related and Real Word - unrelated conditions as they would have the strongest semantic connections. For the novel words, we hypothesized that the Flatness Harmonious words would lead to N400 signals similar to Real Turkish words which means there will be an more negative going mean amplitude for the Flatness Harmonious - unrelated word pairs compared to Flatness Harmonious - related pairs. On the other hand, we expect participants to have a difficulty on semantic consolidation for the Flatness Disharmonious words which would lead to non-significant differences between Flatness Disharmonious - related and Flatness Disharmonious - unrelated condition.

However, if the recalling of the meaning of the words is a more controlled process in the early phases of learning, the neural activity pattern to novel words would be in the LPC time window. According to LPC literature, in the early phases of learning and the semantic consolidation is not strong enough, a higher LPC amplitude is seen to word compared to unrelated words (Bakker et al., 2015; Lei et al., 2022). As the LPC is a signal associated with early learning phase, we do not expect to see a different between Real Word - related and Real Word - unrelated conditions in the LPC time window. Moreover, we do not expect to see a difference between Flatness Harmonious - related and Flatness Harmonious - unrelated words because our expectation is that harmonious words would be more similar to real words. However, we expect Flatness Disharmonious - related words to have a larger amplitude LPC compared to Flatness Disharmonious - unrelated words as Disharmonious words are the deviants than the norm.

Chapter 2

Experiment 1

2.1 Methods

Participants

82 native Turkish speakers (TR) and 39 speakers of other languages (Other) were recruited via email announcement systems and online announcements. The age range was 18-30 (M= 20.24, SD= 2.17, 36 female) for Turkish speakers and 18-57 (M= 29.44, SD= 11.6, 21 female) for the speakers of other languages.

From other-language speakers, participants who had Turkish knowledge and/or had a native language with vowel harmony (n=5, 1 Kazakh, 2 Uyghur, 2 L2 Turkish) were excluded from the analysis (see Table 2.1 for the first language of participants.) In addition, a participant from the native Turkish group was also excluded for repeating the experiment twice. In total, data from 81 participants of native Turkish speakers and 34 participants of other-language speakers with no Turkish knowledge was included in the final results.

Participants who were taking the GE250 course at Bilkent University received course credits for their participation.

Table 2.1: First Language of participants

Language	Number
Arabic	3
Brazilian Portuguese	1
Dutch	1
English	5
German	6
Hebrew	1
Italian	2
Persian	2
Polish	3
Russian	1
Spanish	3
Taiwanese and Mandarin	1
Urdu	3

Materials

The stimuli in the final experiment consisted of a total of 24 auditory novel pseudowords; this included 8 novel pseudowords per condition and 8 familiar object or animal pictures.

Audio Stimuli

For the creation of the novel pseudowords, 36 CVCV Turkish pseudowords were adapted from Altan (2011). All words were not real words in both Turkish and English and they obeyed Turkish phonetic rules. Consonants were drawn from [b], [c], [ç], [d], [f], [g], [h], [k], [m], [n], [p], [s], [ş], [t], [v], [y], and [z]. Each voiceless consonant ([ç], [f], [h], [k], [p], [s], [ş], [t]) was paired with a voiced consonant ([b], [d], [g], [m], [n], [v], [y], [z]) for each word. As in Altan (2011), only voiceless consonants were used in the word-initial position because the production of voiceless consonants between two vowels results in those consonants becoming voiced according to Turkish phonological rules (see consonant alternation in Turkish in Goksel and Kerslake 2011). Voiced consonants were not paired with their voiceless counterparts; therefore, voiceless-voiced counterpart pairs were not used in the same word. This resulted in 8 consonant pairs where each consonant was used only once.

Eight Turkish vowels were divided into groups of four (e,i,o,u and a,ı,ö,ü) where each group included two vowels for back/front, round/unround and high/low characteristics. Since there are two different vowel harmony rules in Turkish, Backness and Flatness, the words in the experiment were designed to include both harmonious and disharmonious versions for each harmony rule. This led to four groups of vowel harmony conditions for the words: (1) harmonious for both flatness and backness harmony rules, (2) harmonious for flatness and disharmonious for backness harmony rules, (3) disharmonious for flatness, and harmonious for backness rules, (4) disharmonious for both flatness harmony and backness harmony rules. Each consonant pair was randomly assigned to one of the vowel groups, and the different vowel harmony conditions for that consonant pair were created with the vowels from the assigned vowel group. Accordingly, a total of 32 words were prepared; however, only 24 words were ultimately used due to repeating some words in two experimental conditions based on the design of the controlled vowel harmony rule (see Design section below for more details).

The stimuli were recorded by a female Turkish native speaker in a silent room in Bilkent University Psychology labs. Noise reduction and amplification of the sounds were used on the recorded words by using Audacity software (Audacity Team, 2021). The intensity was equalized to 70 dB in PRAAT software (Boersma and van Heuven, 2001).

Picture stimuli

Eight familiar object or animal pictures were used as the referent pictures for the novel words. Each referent was selected from a different semantic category from the Turkish adaptation of category norms, where participants were asked to write words from given semantic category and written answers were listed according to how many participants gave the same answer (Cinan, 2001; Battig and Montague, 1969). The 9 most frequently answered 3-8 letter words from each category were listed, and the pictures of the least frequent answers were taken from the Bank Of Standardized Stimuli (BOSS) by Brodeur et al. (2010). In some categories, the second least frequent answers were also used because the corpus either did not have the picture of the least frequent object or the least frequent object was not clearly separable from the object of the same category (e.g., in Turkish, sour cherry vs cherry). Each consonant pair was randomly paired with one of the pictures and that word-referent pair held constant in all conditions. See appendix A for the referent pictures and paired consonants.

Design

The experiment had a 2 (Language Group: Turkish speakers, non-Turkish speakers) x 2 (Flatness: Disharmonious, Harmonious) x 2 (Backness: Disharmonious, Harmonious) design. There were four vowel harmony conditions as a between-subject factor, Flatness Harmonious (FlatHarm), Flatness Disharmonious (FlatDisharm), Backness Harmonious (BackHarm), and Backness Disharmonious (BackDisharm). All the words within that condition had a consistent harmony pattern as named in the condition. For example, all the words within the FlatHarm condition were harmonious for Flatness harmony. Moreover, both Backness and Flatness harmony rules were controlled in the conditions as a within-subject factor. Therefore, if the participant were assigned to the FlatHarm condition, they would hear only the words that had the Flatness harmony pattern, however, for the Backness harmony rule, half of the words they would hear would have the Backness harmony pattern while the other half had the Backness Disharmony pattern. This was the case for all the harmony conditions, so for whichever vowel harmony rule the condition has a constant pattern (Backness harmony or Flatness harmony), there is always a within condition factor for the other harmony rule (e.g. Flatness conditions have both Backness harmonious and Backness disharmonious words and Backness conditions have words both Flatness harmonious and Flatness disharmonious words). From now on, this thesis will use the following labels for the conditions based on the between-participant VH rule: BackHarm, BackDisharm, FlatHarm, FlatDisharm; participants were assigned to one of these four conditions. See Table 2.2 for the stimuli and their condition name.

Each condition had 8 words. These were selected based on the phonological characteristics and constraints of the condition, such that certain words were repeated in two different conditions to ensure that comparisons across conditions were based on the vowel harmony manipulation and not differences in consonant contrasts (e.g., /pime/ was a word in both the Flatness and Backness Harmonious conditions because it is harmonious for both of those characteristics).

Table 2.2: Novel pseudowords used in the experiment

Flatness Harmonious		Flatness Disharmonious	
Backness Harmonious	Backness Disharmonious	Backness Harmonious	Backness Disharmonious
pime	süba	pumo	sıbö
fiye	küva	fuyo	kıvö
şıda	çuge	şüdö	çigo
tıza	hune	tüzö	hino
Backness Harmonious		**Backness Disharmonious**	
Flatness Harmonious	Flatness Disharmonious	Flatness Harmonious	Flatness Disharmonious
pime	çugo	pume	sıbö
fiye	huno	fuye	kıvö
şıda	küvö	tüza	çigo
tıza	sübö	şüda	hino

Procedure

The participants completed an online CSWL experiment which consisted of a training phase and a recognition test phase. The procedure was adopted from the training and recognition test phases of Angwin et al. (2022). The experiment was coded in OpenSesame 3.3.10 (Mathôt et al., 2012) and distributed through JATOS (Lange et al., 2015). Participants ran the experiment on their web browser on a computer. Figure 2.1 illustrates the process of Experiment 1.

Training

The experiment started with the consent page (approved by Bilkent University Ethics Committee) and after the participants consented, they proceeded to the training phase. At the beginning of the training, participants were instructed to listen to the words through headphones while looking at the pictures on the screen. No other instructions about training or information about recognition test were presented.

There were 48 trials in the training and every word-picture pair was presented 8 times. Each word-picture pair was presented with every other pair at least once. The presentation of the auditory words and picture referents was counterbalanced in a way that each image was presented on the right and left sides an equal amount of times. The audio presentation order was counterbalanced according to the referent picture location: when a referent was shown on the left side of the screen, the auditory word corresponding to that referent was presented first for half the trials and second for the other half of the trials. This type of counterbalancing was made to obstruct explicit indication of which auditory word corresponded to which picture (Yu and Smith, 2007; Escudero et al., 2016; Mulak et al., 2019; Angwin et al., 2022).

Each trial in the training phase started with a black fixation point on a white background for 1000ms. After 1000 ms the fixation point disappeared and two referent pictures were presented on the right and left sides of the screen. The first auditory pseudoword was presented 500 ms after the onset of the pictures. The audio for the other picture played 3000 ms after the first word. The pictures remained on the screen for a total of 6500 ms. The trial ended with a blank screen for 1500 ms before the start of the next trial.

Recognition test

Immediately after the training phase ended, the instructions for the recognition test were given. Participants were instructed to listen to one auditorily-presented word while two possible referents were on the screen. They were asked to indicate the correct corresponding picture for the word by pressing the letter ‘A’ on the keyboard if they thought the correct one was on the left or the letter ‘L’ if they thought the correct one was on the right.

There were a total of 32 trials in the recognition test. Each word was served as a target four times and as a distractor four times. Each word was presented against every other word once: for half of those trials, that word was the target, for the other half, that word was the distractor. Therefore, each word was presented with the referent of 2 harmonious and 2 disharmonious words, one for each vowel

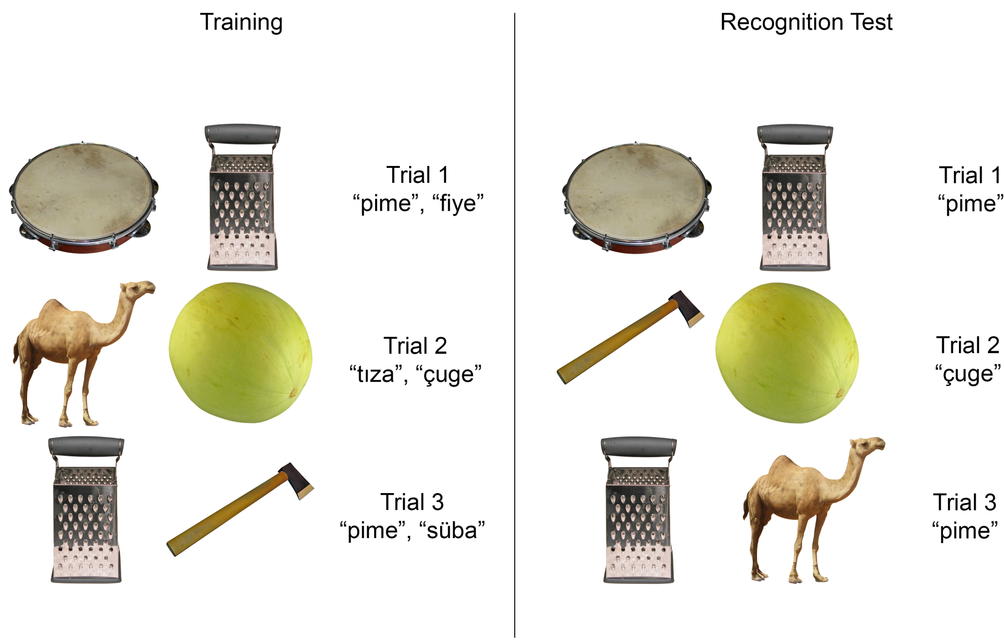

Figure 2.1: An process illustration of Experiment 1

group. This meant that words belonging to the same VH pattern with the same vowels were presented against each other twice, once as the target and once as the distractor (e.g., /pime/ vs /fiye/). The location of the target referent picture was counterbalanced across trials.

The trials started with a fixation point of 1000ms. After the fixation point, two referent pictures were presented on the right and left sides of the screen. With the onset of the picture screen, the target word was presented auditorily. The pictures stayed on the screen either until an answer key was pressed or for 5 seconds. A white screen for 1500 ms was presented before the next trial started. No feedback for accuracy was given.

2.2 Results

Overall, both native Turkish speakers and speakers of other languages were successful on the Recognition test with above-chance performance (87% accuracy for

TR , 90% accuracy for Other).

Because of the characteristics of Turkish vowel harmony rules, the analysis of the experiment was done by comparing harmonious and disharmonious conditions of each vowel harmony rule instead of comparing four groups. Therefore the analysis was done separately for Backness Harmonious/Disharmonious conditions and Flatness Harmonious/Disharmonious conditions separately.

For the Flatness harmony between subject conditions, a logistic mixed model (bobyqa optimizer) was fitted to predict accuracy with fixed effects of Backness Harmony, Flatness Harmony, and Language group and a random effect of participants. The model did not find a significant effect of any of the factors or interactions. A linear fixed model (bobyqa optimizer) was fitted to predict reaction time with fixed effects of Backness Harmony, Flatness Harmony, and Language group and a random effect of participants. The model did not find significant effects of interactions of any factors.

For the Backness harmony conditions, a logistic mixed model (bobyqa optimizer) to predict accuracy with fixed effects of Backness Harmony, Flatness Harmony, and Language group and a random effect of participants was fitted. None of the factors or the interaction were significant. A linear mixed model (bobyqa optimizer) to predict reaction time was fitted with the same predictors. The effects of predictors were not significant.

Table 2.3: Mean accuracy of the participants in Experiment 1

	TR		Other	
	Backness Harmonious	Backness Disharmonious	Backness Harmonious	Backness Disharmonious
Flatness Harmonious	86.9%	87.4%	91.5%	96.3%
Flatness Disharmonious	88.7%	88.6%	84.6%	89.2%

2.3 Discussion

In the first experiment, we were interested in investigating the effects of vowel harmony types on word learning on both native speakers of Turkish and speakers of other languages. Participants were expected to learn eight words via 2 words - 2 referents CSWL paradigm. We hypothesized to find a lower accuracy rate for the disharmonious words for the native Turkish speakers while we did not expect such a difference for the speakers of other languages.

The participants were highly successful in the recognition test by correctly identifying the correct word-referent pairs. The result of the experiment revealed that both native Turkish speakers and speakers of other languages were able to learn novel words via CSWL, which was in line with our hypothesis. However, the influence of vowel harmony conditions was not significant for neither the accuracy nor the reaction time.

Although there were no significant effects of vowel harmony conditions on learning for either Language group, the high performance of the participants might have hindered the effect. Even though there is no effect, numerically, Other speakers seem to have a higher accuracy in the Flatness Harmonious conditions compared to the Flatness Disharmonious condition, while it is around the same level for the Turkish participants. We believe that might be due to having a higher number of ü and ö sounds in the Flatness disharmony condition, which are Turkish vowels that may not exist in their own languages. That might make it more difficult for the Other group to differentiate sounds in line with the PAM and SLM models which propose that as the novel sounds are different than the sound of the L1 it is more difficult to differentiate (Best and Tyler, 2007; Flege, 1995). However, since all the words were successfully learned in this experiment without any difference between groups, we can only speculate based on mean differences and results should be interpreted cautiously.

The high level of accuracy shows that the design was relatively easy for the participants. The ease of the task might have led to non-significant results for vowel harmony. One way to measure such an effect might be to make the experiment more ambiguous, either by increasing the ambiguity or the number of words to learn.

The biggest limitation of this experiment was the imbalance between the Language groups. The number of participants in the Other group is less than half of the TR group. Unfortunately, we were not able to reach speakers of other languages who do not have any experience with vowel harmonic languages as successfully, which may have created a bias in the results and not shown a true difference in vowel harmony processing.

Chapter 3

Experiment 2

3.1 Methods

Participants

55 native Turkish speakers were recruited via email announcements (mean age: 22.07, SD: 2.49, 22 males). All participants were Turkish - English bilinguals who started learning English as a second language in primary school. Data from 1 participant was excluded due to participating in Experiment 1.

Before starting to the experiment, participants filled out EEG Pre-Screening Form (see appendix D) and Language History Questionnaire (see appendix E).

Participants received either 10 points for GE250 course or 50 TL (approx 3.1 €) for their participation.

Stimuli

Stimuli from two conditions of the first experiment, Backness harmony and Backness Disharmony, were used in the training and recognition test phases of the EEG experiment. This resulted in a total of 16 auditory novel pseudowords such that there were 8 words per condition. The audio words were paired with the same picture referents from Experiment 1, such that the referent pictures were the least or the second least frequent answer from eight different semantic categories from the Turkish adaptation of category norms (Battig and Montague, 1969; Cinan, 2001).

The eight most frequent Turkish words (limited to 1 to 3 syllables) from the same semantic category as the referents in Experiment 1 were selected. The selected real words were paired with pseudowords in the SJT. These real words were presented in the orthographic format in the SJT after the auditorily-presented pseudoword (see Procedure for more details).

In addition to the pseudowords, 8 semantic categories that were not selected for the pseudoword categories were selected to construct the Real Word trials in the SJT from the same Turkish adaptation of category norms (Battig and Montague, 1969; Cinan, 2001). From each category, the least frequent Turkish word of the 9 most frequently answered 1-3 syllable words was selected for auditory presentation. For the Real Word trials, the remaining more frequent 8 Turkish words were paired with real Turkish auditory words and presented orthographically after the auditorily-presented real word. See appendix B for the list of auditorily presented words and appendix C the list of orthographically presented words.

All the pseudowords used in Experiments 1 and 2 and the auditorily presented Turkish real words were recorded by the same native Turkish speaker on the same day.

Procedure

The experiment's procedure was the same as (Angwin et al., 2022) which consisted of three phases; a training phase, a recognition test phase, and a semantic judgment phase in order. Figure 3.1 illustrates the whole process of Experiment 2. The training and recognition task phases were the same as in Experiment 1. The experiment was done in a single session with short breaks between the three

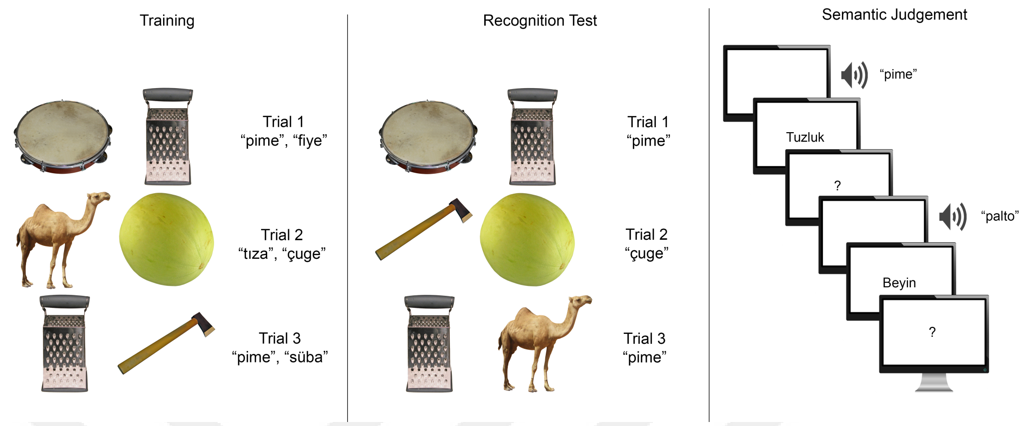

Figure 3.1: An process illustration of Experiment 2

phases. The experiment was coded on MATLAB version 9.1 (The MathWorks Inc., 2022) software with Psychtoolbox 3.0 (Brainard, 1997). In the beginning of the experiment, participants were instructed to try not to blink during the trials and blink when the blank screen appeared.

Training Phase In the training phase, there were a total of 48 trials with the presentation of each word 12 times. To summarize, 2 referent pictures were presented on the two sides of the screen for 6500 ms. 500ms after the presentation of the pictures, the auditory pseudoword of one picture was presented through headphones. The second auditory pseudoword was presented 3000ms after the first one. The location of the pictures and the audio presentation order were counterbalanced within trials and between participants. The trials were presented in a pseudorandom order where the same audio was not presented more than three times in a row.

Recognition Test

Immediately following the training phase, participants started the recognition test phase. Each word was presented four times as a target and four times as a distractor. As in Experiment 1, participants were asked to identify the correct picture referent for the presented audio word by pressing the letter "A" on the keyboard to select the picture on the left and pressing the letter "L" to select the picture on the right. The presentation of the trials was pseudorandomized in a way that the same audio was not presented more than twice in a row.

Semantic Judgment Task

Participants saw the instructions for the SJT on the screen after the recognition test and were asked if they needed a break before starting the SJT. Instructions were also explained by the research assistant.

The Semantic Judgment phase started with four practice trials in the presence of a research assistant. Two auditory Turkish words and four orthographic Turkish words that were not used in the experiment were used in the practice trials. In the practice trials, both auditory Turkish words were paired with one orthographic word from the same semantic category and one from a different semantic category. Trials were counterbalanced in a way such that the same auditory word trials were presented one after the other to emphasize the difference.

After the practice trials, the instructions were given once more and the participants started the SJT. A total of 128 trials were presented on the SJT. 64 of those trials were the presentation of novel pseudowords that the participants learned in the training phase. Each word was presented 8 times: in four of the presentations, each word was paired with another word from the same category, and in the other four trials, they were paired with the words from the categories of other novel words. Since the novel words can be FlatHarm or FlatDisharm, this pairing ends up with four conditions for novel words (16 trials each); Harmonious Novel Word - Same category, Harmonious Novel Word - Different category, Disharmonious Novel Word - Same category and Disharmonious Novel Word - Different category.

The remaining 64 trials were with Real Turkish words which belonged to semantic categories different from the novel words. As for the novel words, Real words were also paired with the word from the same category four times, and paired with the words from the other Real words categories four times. This pairing created Real Word - Same category (32 trials) and Real Word - Different category (32 trials) conditions.

The order of the trials was pseudorandomized in a way that no more than four trials from the same type (Novel/Real word or Same/Different category) were repeated.

The trials were started with a fixation point (+) for 1000 ms. When the fixation point disappeared, an audio word was presented through headphones. The blank screen stayed for 1000 ms. Following the blank screen, the paired orthographic Turkish word was presented on the screen for 1000 ms, which was followed by another blank screen for 1000 ms. After that, a question mark (?) was presented on the screen to indicate to the participants that they should respond whether the presented auditory and written words were from the same semantic category or not. Participants were instructed to press the key 1 if the audio word and the written word were from the same semantic category and press the key 0 if the audio word and the written word were from different semantic categories. The trial ended with a blank screen for 1500 ms before the next trial started.

EEG recording and pre-processing

Electroencephalogram (EEG) data were recorded via 64 channel MR-compatible system (Brain Products GmbH, Gilching, Germany) with Ag/AgCl passive electrodes mounted onto the cap according to the international 10/20 system. 63 of the channels were on the cap while another electrode was placed on the back of the participants for electrocardiogram (ECG) recording. Two electrodes on the scalp, AFz and FCz, were used as ground and reference electrodes respectively. The electrode impedances were held below 10kΩ during the experimental session. The EEG data were recorded at a 1kHz sampling rate.

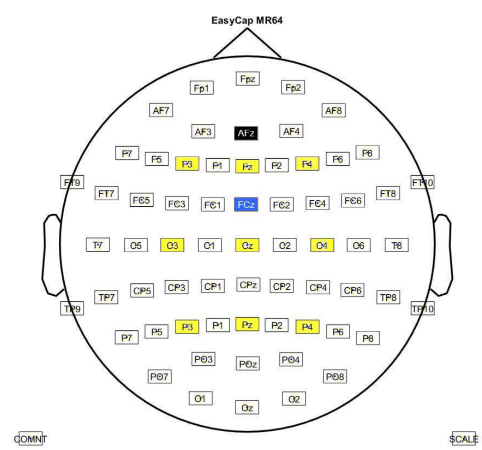

Figure 3.2: The position distribution of the selected electrodes on the EEG cap. The figure is created on MATLAB (The MathWorks Inc., 2022) with FieldTrip Toolbox (Oostenveld et al., 2011)

.

The raw data were preprocessed by using EEGlab (Delorme and Makeig, 2004) and ERPlab (Lopez-Calderon and Luck, 2014) toolboxes and MATLAB scripts. ECG channel was removed from the data. The data were re-referenced offline to the common average and the sampling rate was downsampled to 512Hz. The continuous data were divided into epochs of 200 msec prestimulus onset and 1000 msec after the onset of the stimulus. The onset point of the epoch was the presentation of the written word in the SJT. A 0.1Hz high pass and 30 Hz lowpass bandpass filters were applied to the epochs. Eye movements and noisy channels were detected and corrected by Independent component analysis (ICA) using EEGlab's runica function. ERPlab's Moving Window Peak-to-Peak function with a 100 uV voltage threshold was applied for artifact rejection. Participants (n = 7) with 40% rejected trials overall or 50% rejected trials in more than two conditions were excluded from further analysis.

The EEG analysis was focused upon 9 electrode channels across the scalp (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4). Figure 3.2 shows the position distribution of the selected electrodes. For the N400 ERP, mean amplitudes from 300 - 500 ms time window, and for the LPC, mean amplitudes from 500 - 700 ms time window extracted for the analysis.

3.2 Results

3.2.1 Behavioral Results

One participant who participated in the first experiment was excluded from the analysis for knowing the experimental paradigm. From the remaining data, the trials where the reaction time was either zero or 2sd above or below the mean were excluded from both the Recognition test and the SJT for the analysis (Ratcliff, 1993). This resulted in removing 107 (6.36%) trials from the Recognition test and 408 (6.12%) trials from the SJT.

Recognition Test Accuracy

In the recognition test, participants answered 90% of the trials correctly; 92% for the Backness Harmony condition and 89% for the Backness Disharmony condition. A mixed-effect model analysis (using bobyqa optimizer) was conducted to predict accuracy with Backness Harmony and Flatness Harmony. The reference was selected Backness Harmonious for the Backness Harmony factor and FlatHarm for the Flatness Harmony factor.

None of the factors nor the interaction was significant.

Recognition Test Reaction Time

A mixed effect model analysis to predict reaction time (using bobyqa optimizer) is conducted to see the effects of Backness Harmony and Flatness Harmony, considering participants as a random effect, on the reaction time. The reference was selected Backness Harmonious for the Backness Harmony factor and FlatHarm for the Flatness Harmony factor.

According to the model, the effect of Backness Harmony was not significant.

There were no main effects of the Backness Harmony factor, and there was no interaction between the Backness Harmony and the Flatness Harmony factors.

Considering Backness Harmonious and FlatHarm conditions as the baseline, the effect of Flatness Harmony was significant and positive according to the model (β = 7.54, 95% CI [5.36, 9.73],$t(1569)$ = 6.77,p <.001; Std. β = 0.11, 95% CI [0.08, 0.15]). The pairwise test showed that participants were faster at the Flatness Disharmony (M = 124 ms, SE = 7.41) condition than the Flatness Harmony condition (M = 139 ms, SE = 7.33).

The interaction of the Backness Harmony and the Flatness Harmony was not significant for the reaction time.

Semantic Judgment Task Accuracy

In the Semantic Judgement task, participants showed a successful performance with 87% accuracy, with 85.4% accuracy for the Backness Disharmony condition and 90.7% accuracy for the Backness Harmony condition.

Three separate model-fittings were conducted for accuracy. The first model, Model 1, investigated fixed effects of Word Type (Real/Novel), Category Congruency (related/unrelated) and Backness Harmony (Backness Harmonious/ Backness Disharmonious). Model 2 investigated the fixed effects of Flatness Harmony (Real/FlatHarm, FlatDisharm), Category Congruency (related/unrelated) and Backness Harmony (Backness Harmonious/ Backness Disharmonious). And lastly, the Model 3 combined Flatness Harmony and Category Congruency factors investigated the fixed effects of Flatness Harmony - Category Congruency (FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated, Real Word - related, and Real Word - unrelated) and Backness Harmony (Backness Harmonious/ Backness Disharmonious) to have a similar model fitting to EEG analysis. Since EEG triggers were sent for combined Flatness Harmony - Category Congruency factor, EEG analysis did not have separate Flatness Harmony and Category Congruency factors.

Model 1 was fitted by considering the effects of Word Type (Real/Novel), Category Congruency (related/unrelated), and Backness Harmony (Backness Harmonious/Backness Disharmonious), while participant was included as a random effect. For each factor, the reference point was selected as follows: Real words for Word Type (vs. Novel), Related for Category Congruency (vs. unrelated), and Backness Harmonious for the Backness Harmony factor (vs. Backness Disharmonious).

According to the model, the effect of Backness Harmony was not significant.

The effect of Word Type was significant (β = -1.44, 95% CI [-1.62, -1.25],p <.001; Std. β = -1.44, 95% CI [-1.62, -1.25]). The Bonferroni-corrected pairwise comparisons showed that participants were more accurate on Real words (M =

95.9%, SD = 0.005) compared to Novel words (M = 84.9%, SD = 0.016) (p <.001).

Category Congruency also had a significant effect on accuracy (β = -0.24, 95% CI [-0.39, -0.08],p = 0.004; Std. β = -0.24, 95% CI [-0.39, -0.08]). Follow-up Bonferroni-corrected pairwise comparisons showed that participants were more accurate on unrelated pairs (M= 93.6%, SE = 0.008) compared to the related pairs (M = 90%, SE = 0.012) (p <.001).

There was no significant interaction.

In order to investigate the fixed effect of the flatness harmony condition, the data were fitted to another mixed-effects model (Model 2), where the baseline for the flatness harmony factor was Real words, for Congruency was related, and for the Backness harmony was Backness harmonious. Participant was included as a random effect.

According to this second model, there was effect of Flatness harmony was significant for both FlatHarm (β = -1.57, 95% CI [-1.78, -1.37],p <.001; Std. β = -1.57, 95% CI [-1.78, -1.37]) and FlatDisharm (β = -1.27, 95% CI [-1.49, -1.06],p <.001; Std. β = -1.27, 95% CI [-1.49, -1.06]) conditions. To further investigate the Flatness Harmony factor, Bonferroni-corrected pairwise comparisons conducted for the Flatness Harmony factor. According to the model and pairwise comparisons, participants were more accurate on Real Word trials (M = 95.9%, SE = 0.005) than FlatHarm (M = 83%, SE = 0.019) and FlatDisharm (M = 86%, SE = 0.016) trials. The pairwise comparisons showed a difference between FlatHarm and FlatDisharm conditions where participants answered more accurately in FlatDisharm trials compared to FlatHarm trials (p = .009).

The model also showed an significant effect of Category Congruency (β = -0.24, 95% CI [-0.39, -0.08],p = 0.004; Std. β = -0.24, 95% CI [-0.39, -0.08]) where participants were more accurate on the unrelated trials (M= 92.2%, SE = 0.009) compared to related trials (M= 87.7%, SE = 0.014) (p <.001).

The model also showed an interaction between Flatness Harmony and Backness Harmony for the Backness Disharmonious x FlatDisharm conditions (β = 0.22, 95% CI [2.52e-04, 0.43],p = 0.050; Std. β = 0.22, 95% CI [2.52e-04, 0.43]). A Bonferroni-corrected pairwise comparison was run to investigate the interaction. According to the comparison, the Real Word - Backness Harmonious (M = 96.4%, SE = 0.007) condition was significantly more accurate than FlatHarm - Backness Harmonious (M = 86.9%, SE = 0.022) (p <.001), FlatHarm - Backness Disharmonious (M = 78.3%, SE = 0.032) (p <.001), FlatDisharm - Backness Harmonious (M = 90.4%, SE = 0.017) (p <.001) and FlatDisharm - Backness Disharmonious (M = 82.3%, SE = 0.027) (p <.001) conditions. The comparison also indicated that Real Word - Backness Disharmonious (M = 95.4%, SE = 0.008) condition was significantly more accurate than FlatHarm - Backness Harmonious (p <.001), FlatHarm - Backness Disharmonious (p <.001) and FlatDisharm - Backness Disharmonious (p <.001) conditions. There was also a significant difference between FlatHarm - Backness Harmonious and FlatDisharm - Backness Harmonious conditions, where FlatDisharm - Backness Harmonious was more accurate (p = .009).

Model 3 was fitted to predict the accuracy with Flatness Harmony - Category Congruency (FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated, Real Word - related, and Real Word - unrelated) and Backness Harmony (Backness Harmonious/ Backness Dishamronious) by combining Flatness Harmony - Category Congruency similar to EEG analysis. The reference was Real Word - related for Flatness Harmony - Category Congruency factor and Backness Harmonious for the Backness Harmony factor.

The effect of the Backness Harmony factor was not significant according to the model.

The model showed a significant effect of Flatness Harmony - Category Congruency for FlatHarm-related (β = -1.53, 95% CI [-1.80, -1.25],p <.001; Std. β = -1.53, 95% CI [-1.80, -1.25]), FlatHarm-unrelated (β = -1.15, 95% CI [-1.43, -0.87],p <.001; Std. β = -1.15, 95% CI [-1.43, -0.87]), FlatDisharm-related (β = -1.36, 95% CI [-1.64, -1.09],p <.001; Std. β = -1.36, 95% CI[-1.64, -1.09]), FlatDisharm-unrelated (β = -0.71, 95% CI [-1.02, -0.41],p <.001; Std. β = -0.71, 95% CI[-1.02, -0.41]), and Real Word - unrelated (β = 0.47, 95% CI [0.15, 0.79],p = 0.004; Std. β = 0.47, 95% CI [0.15, 0.79]).

To further investigate the effect, a Bonferroni-corrected pairwise comparison was conducted. The results of the pairwise test implied a significant difference between Real Word - related (M = 94%, SE = 0.007) condition and FlatHarm-related (M = 80.2%, SE = 0.24)(p <.001), FlatHarm-unrelated (M = 85.5%, SE = 0.019) (p <.001), FlatDisharm-related (M = 82.7%, SE = 0.021) (p <.001), and FlatDisharm-unrelated (M = 90.2%, SE = 0.014)(p = 0.0001) trials where Real word - related trials had a higher accuracy than all novel word conditions. The pairwise comparison also showed that Real Words - unrelated trials had a higher accuracy than FlatHarm-related (p <.001), FlatHarm-unrelated (p <.001), FlatDisharm-related (p <.001), and FlatDisharm-unrelated (p <0.001) trials. There was a marginally significant difference between Real Word - related and Real Word - unrelated conditions (p = 0.052) where Real Word - unrelated conditions had a higher accuracy. Among the novel word conditions, the accuracy for FlatDisharm - unrelated was significantly higher compared to FlatDisharm - related (p = 0.0002) and FlatHarm-related (p <.001).

There was no significant interaction between Flatness Harmony - Category Congruency and Backness Harmony.

Semantic Judgment Task Reaction Time

Similar to accuracy analysis, three separate linear mixed effect models were fitted for reaction time, although reaction time analysis should be interpreted with caution as there was a 1-second delay before answering in the SJT.

To investigate the effect of Real word Novel word difference, a linear mixed effect model (Model 1) was fitted with the reference for the flatness harmony factor was Real words, for Congruency was related, and for the Backness harmony was Backness harmonious, and participant was included as a random effect.

Within the model, the effect of Backness Harmony was not significant. According to the model, the effect of Word Type was significant (β = 155.57, 95% CI [133.09, 178.05],$t(6238)$ = 13.57,p <.001; Std. β = 0.32, 95% CI [0.27, 0.36]). Follow up pairwise analysis indicatated that participants were faster at Real Word trials (M = 544 ms, SE = 25.7) compared to Novel Word trials (M = 700 ms, SE = 25.7) (p <.001). Category Congruency also showed a significant effect for the model (β = -25.15, 95% CI [-40.87, -9.43],$t(6238)$ = -3.14,p = 0.002; Std. β = -0.05, 95% CI[-0.08, -0.02]). According to that, participant were faster at answering related trials (M = 598 ms, SE= 25.7) compared to unrelated trials (M = 646 ms, SE=25.7). There were no other main effects and no interactions.

To see the predictive effect of Flatness harmony, another linear mixed-effects model was fit (Model 2), where the baseline for the flatness harmony factor was Real words, for Congruency was related, and for the Backness harmony was Backness harmonious, and the participants were a random effect.

There were no effect of Backness Harmony on the reaction time. According to the model, there was an effect of both FlatHarm (β = 178.77, 95% CI [151.17, 206.37],$t(6234)$ = 12.70,p <.001; Std. β = 0.36, 95% CI [0.31, 0.42]) and FlatDisharm (β = 132.26, 95% CI [104.68, 159.85],$t(6234)$ = 9.40,p <.001; Std. β = 0.27, 95% CI [0.21, 0.33]) levels. To further investigate the effect, Bonferroni-corrected pairwise comparisons were conducted. According to the pairwise comparisons, participants response to Real Word trials (M = 544ms, SE = 25.7) significantly faster than FlatHarm (M = 723 ms, SE = 27)(p <.001) and FlatDisharm (M = 677 ms, SE = 27) (p <001) trials. There was also a significant difference between FlatHarm and FlatDisharm conditions (p = .013).

The effect of Category Congruency was also significant (β = -25.14, 95% CI [-40.85, -9.44],$t(6234)$ = -3.14,p = 0.002; Std. β = -0.05, 95% CI [-0.08, -0.02]). Accordingly, participant were faster at related conditions (M = 624 ms, SE = 25.9) compared to unrelated conditions (M = 672 ms, SE = 25.8).

There was a significant interaction of Backness Harmonious x FlatDisharm trials in the model 2 (β = -37.79, 95% CI [-65.38, -10.20],$t(6234)$ = -2.69,p = 0.007;

Std. β = -0.08, 95% CI [-0.13, -0.02]). To further investigate the interaction, Bonferroni-corrected pairwise comparisons were conducted. The results of the pairwise test implied that participants were faster at Real Word - Backness Harmonious condition (M = 554 ms, SE = 36.3) compared to FlatHarm - Backness Harmonious (M = 731 ms, SE = 38.2)(p <.001), FlatHarm - Backness Disharmonious (M = 715ms, SE = 38.2)(p = .03) and FlatDisharm - Backness Disharmonious (M = 705ms, SE = 38.1)(p <.001) conditions. Real Word- Backness Disharmonious condition (M = 534 ms, SE = 36.3) also lead to faster response than FlatHarm - Backness Harmonious (p = .002), FlatHarm - Backness Disharmonious (p <.001) and FlatDisharm - Backness Disharmonious (M = 649 ms, SE = 38.2)(p <.001) conditions. There was also a significant difference between FlatDisharm - Backness Harmonious and FlatHarm - Backness Harmonious conditions (p = .005), where FlatDisharm - Backness Harmonious condition led to faster responses.

Another mixed effects model was fitted to predict the reaction time with Flatness Harmony - Category Congruency (FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated, Real Word - related, and Real Word - unrelated) and Backness Harmony (Backness Harmonious/ Backness Dishamronious) by combining Flatness Harmony - Category Congruency similar to EEG analysis. The reference was Real Word - related for Flatness Harmony - Category Congruency factor and Backness Harmonious for the Backness Harmony factor.

The effect of Backness Harmony on the model was not significant. However, there was an effect of Flatness Harmony - Category Congruency factor on FlatHarm - related (β = 188.49, 95% CI [149.16, 227.81],$t(6234)$ = 9.40,p <.001; Std. β = 0.38, 95% CI [0.30, 0.46]) , FlatHarm - unrelated (β = 219.34, 95% CI [180.57, 258.11],$t(6234)$ = 11.09,p <.001; Std. β = 0.45, 95% CI [0.37, 0.53]), FlatDisharm - related (β = 127.24, 95% CI [88.30, 166.18],$t(6234)$ = 6.41,p <.001; Std. β = 0.26, 95% CI [0.18, 0.34]), FlatDisharm - unrelated (β = 187.58, 95% CI [148.45, 226.70],$t(6234)$ = 9.40,p <.001; Std. β = 0.38, 95% CI [0.30, 0.46]) and Real Word - unrelated (β = 50.29, 95% CI [18.88, 81.70],$t(6234)$ = 3.14,p = 0.002; Std. β = 0.10, 95% CI [0.04, 0.17]). To further investigate the effect, a pairwise comparison was done for the Flatness Harmony - Category Congruency factor. According to the pairwise comparisons, participants were significantly faster in the Real Word- related (M = 519 ms, SE = 26.9) condition compared to FlatHarm - related (M = 708 ms, SE = 26.9)(p <.001), FlatHarm - unrelated (M = 719 ms, SE = 29.3) (p <.001), FlatDisharm - related (M = 646 ms, SE = 29.3)(p <.001), FlatHarm - unrelated (M = 707 ms, SE = 29.4)(p <.001) and Real Word - unrelated (M = 570 ms, SE = 26.9)(p = .025)). Real Word- unrelated condition is also significantly fater than all novel word conditions, FlatHarm - related (p <.001), FlatHarm - unrelated (p <.001), FlatDisharm - related (p = .0016) and FlatDisharm - unrelated (p <.001). The comparison also implied a significant difference between FlatDisharm- related and FlatHarm - unrelated conditions where FlatDisharm- related condition was significantly faster than FlatHarm - unrelated (p = .0009).

According to the model, there was also an interaction effect of Backness Disharmonius x FlatDisharm condition (β = -42.81, 95% CI [-81.93, -3.69],t(6234) = -2.15,p = 0.032; Std. β = -0.09, 95% CI [-0.17, -7.52e-03]). Bonferroni-corrected pairwise comparisons were conducted to investigate the interaction between Backness Harmony and Flatness Harmony factors. See Table 3.1 for the mean and standard deviation values of Backness Harmony and Flatness Harmony interaction and table 3.2 for the full list of significant differences.

Table 3.1: Mean and SE values of Flatness Harmony - Category Congruency interaction for the SJT reaction time.

Backness Harmony	Flatness Harmony - Category Congruency	Mean	SE
Backness Harmonious	Real Word - related	523.41	38.07
Backness Disharmonious	Real Word - related	515.07	38.02
Backness Harmonious	FD - unrelated	668.18	41.64
Backness Disharmonious	FD - unrelated	745.45	41.50
Backness Harmonious	FD - related	629.31	41.56
Backness Disharmonious	FD - related	663.65	41.43
Backness Harmonious	FH - unrelated	771.32	41.33
Backness Disharmonious	FH - unrelated	705.83	41.45
Backness Harmonious	FH - related	690.85	41.69
Backness Disharmonious	FH - related	724.60	41.67
Backness Harmonious	Real Word - unrelated	585.13	38.12
Backness Disharmonious	Real Word - unrelated	553.92	37.94

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious
Note. All values are in milliseconds.

Table 3.2: The interaction table for Flatness Harmony - Category Congruency for the SJT reaction time.

		Mean Difference	SE	z.ratio	p.value
BH - Real Word-related	BD - Real Word-related	8.344	53.802	0.155	1.000
	BH - FD - unrelated	-144.767	28.347	-5.107	<.0001
	BD - FD - unrelated	-222.040	56.316	-3.943	0.005
	BH - FD - related	-105.900	28.182	-3.758	0.011
	BD - FD - related	-140.235	56.263	-2.492	0.837
	BH - FH - unrelated	-247.915	27.904	-8.884	<.0001
	BD - FH - unrelated	-182.421	56.277	-3.241	0.079
	BH - FH - related	-167.438	28.384	-5.899	<.0001
	BD - FH - related	-201.192	56.441	-3.565	0.024
	BH - Real Word-unrelated	-61.722	22.805	-2.706	0.449
	BD - Real Word-unrelated	-30.513	53.744	-0.568	1.000
BD - Real Word-related	BH - FD - unrelated	-153.111	56.389	-2.715	0.437
	BD - FD - unrelated	-230.383	28.099	-8.199	<.0001
	BH - FD - related	-114.244	56.325	-2.028	1.000
	BD - FD - related	-148.579	27.995	-5.307	<.0001
	BH - FH - unrelated	-256.259	56.161	-4.563	0.000
	BD - FH - unrelated	-190.765	28.029	-6.806	<.0001
	BH - FH - related	-175.782	56.423	-3.115	0.121
	BD - FH - related	-209.535	28.352	-7.391	<.0001
	BH - Real Word-unrelated	-70.065	53.838	-1.301	1.000
	BD - Real Word-unrelated	-38.856	22.515	-1.726	1.000
BH - FD - unrelated	BD - FD - unrelated	-77.272	58.792	-1.314	1.000
	BH - FD - related	38.867	32.870	1.182	1.000
	BD - FD - related	4.532	58.741	0.077	1.000
	BH - FH - unrelated	-103.148	32.619	-3.162	0.103
	BD - FH - unrelated	-37.654	58.755	-0.641	1.000
	BH - FH - related	-22.671	33.045	-0.686	1.000
	BD - FH - related	-56.424	58.912	-0.958	1.000
	BH - Real Word-unrelated	83.046	28.405	2.924	0.228
	BD - Real Word-unrelated	114.254	56.333	2.028	1.000
BD - FD - unrelated	BH - FD - related	116.139	58.732	1.977	1.000
	BD - FD - related	81.804	32.560	2.512	0.792
	BH - FH - unrelated	-25.875	58.574	-0.442	1.000
	BD - FH - unrelated	39.618	32.583	1.216	1.000
	BH - FH - related	54.601	58.825	0.928	1.000
	BD - FH - related	20.848	32.863	0.634	1.000
	BH - Real Word-unrelated	160.318	56.351	2.845	0.293
	BD - Real Word-unrelated	191.527	27.995	6.841	<.0001
BH - FD - related	BD - FD - related	-34.335	58.681	-0.585	1.000
	BH - FH - unrelated	-142.015	32.501	-4.370	0.001
	BD - FH - unrelated	-76.521	58.694	-1.304	1.000
	BH - FH - related	-61.538	32.901	-1.870	1.000
	BD - FH - related	-95.291	58.851	-1.619	1.000
	BH - Real Word-unrelated	44.179	28.228	1.565	1.000
	BD - Real Word-unrelated	75.388	56.270	1.340	1.000
BD - FD - related	BH - FH - unrelated	-107.680	58.523	-1.840	1.000

	BD - FH - unrelated	-42.186	32.494	-1.298	1.000
	BH - FH - related	-27.203	58.774	-0.463	1.000
	BD - FH - related	-60.956	32.776	-1.860	1.000
	BH - Real Word-unrelated	78.514	56.298	1.395	1.000
	BD - Real Word-unrelated	109.722	27.888	3.934	0.006
BH - FH - unrelated	BD - FH - unrelated	65.494	58.536	1.119	1.000
	BH - FH - related	80.477	32.667	2.464	0.908
	BD - FH - related	46.723	58.694	0.796	1.000
	BH - Real Word-unrelated	186.193	27.968	6.657	<.0001
	BD - Real Word-unrelated	217.402	56.105	3.875	0.007
BD - FH - unrelated	BH - FH - related	14.983	58.788	0.255	1.000
	BD - FH - related	-18.770	32.802	-0.572	1.000
	BH - Real Word-unrelated	120.699	56.312	2.143	1.000
	BD - Real Word-unrelated	151.908	27.916	5.442	<.0001
BH - FH - related	BD - FH - related	-33.753	58.944	-0.573	1.000
	BH - Real Word-unrelated	105.717	28.433	3.718	0.013
	BD - Real Word-unrelated	136.925	56.368	2.429	0.999
BD - FH - related	BH - Real Word-unrelated	139.470	56.475	2.470	0.893
	BD - Real Word-unrelated	170.679	28.245	6.043	<.0001
BH - Real Word-unrelated	BD - Real Word-unrelated	31.209	53.780	0.580	1.000

Abbreviations. BH, Backness Harmonious; BD, Backness Disharmonious; FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. Mean Difference and SE values are in millisecond.

Although the fitted models showed significant differences in the reaction times, the results should be interpreted with caution as there was a 1-second delay before responses were given.

Overall, the SJT results showed that participants were more accurate when two words were from different categories. In particular, participants had difficulty in correctly grouping novel words when they were presented with Real words from the same semantic category. For the reaction time, participants were faster for the real words and when the words were from the same semantic category, but there was no interaction between the word type and Category Congruency. Moreover, considering the effect of flatness harmony, participants showed a higher accuracy in the FlatDisharm - unrelated than FlatDisharm - related, while FlatHarm words did not show this result.

3.2.2 EEG measurements

As in the behavioral experiment analysis, the data from the participant who participated in the first experiment was excluded from the EEG analysis. In addition, 6 participants were excluded from the EEG analysis due to excessive eye blinking (three from the Backness Harmony and three from the Backness Disharmony conditions). Therefore, data from 47 participants were included in the final analysis. Figure 3.3 shows the averaged ERPs from -200 ms to 1000 ms from the selected electrodes for each condition.

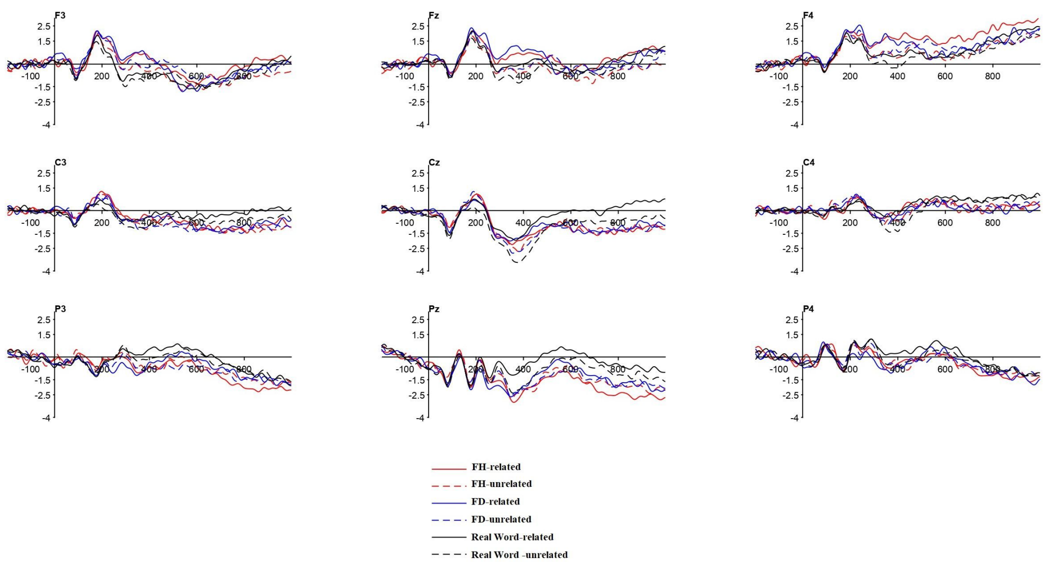

Figure 3.3: Averaged ERP waveforms for the selected electrodes

N400

A repeated measures ANOVA with Flatness Harmony-Category Congruency (6 levels: FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated, Real Word - related, and Real Word - unrelated), anteriority (3 levels: frontal, central, posterior), and lateralization (3 levels: left, midline, right) as within-subject factors and Backness harmony (2 levels: Backness Harmonious, Backness Disharmonious) as a between-subject factor was conducted to examine the N400. Figure 3.4 shows the ERP waveform for Cz electrode with marked N400 time window.

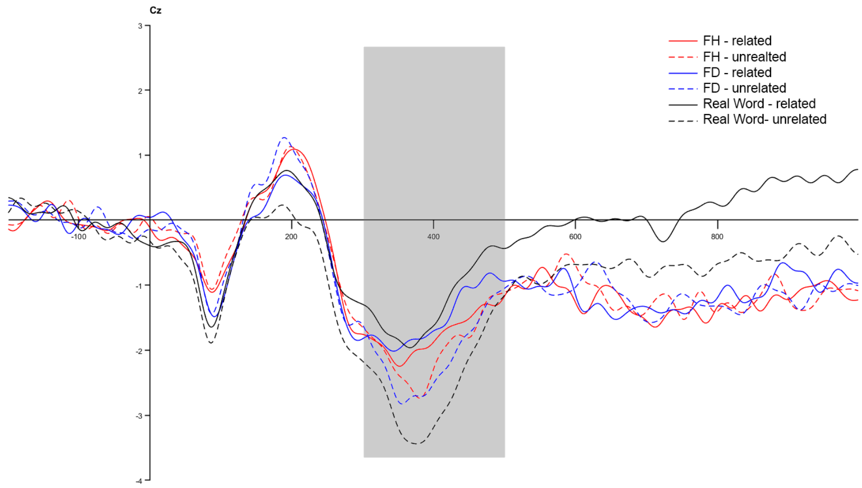

Figure 3.4: Averaged ERP waveforms for Cz electrode, N400 time window is marked

The results revealed a significant main effect of Flatness Harmony-Category Congruency (F(4.143, 190,561)=9.839,p <.001, η^2p = 0.176). See Table 3.3 for the mean amplitude values for conditions. Follow-up Bonferroni-corrected pairwise comparisons demonstrated that the N400 amplitude for Real Word - unrelated (M = 0.83 mV, SD = 2.28) was more negative than FlatHarm-related (M = -0.32 mV, SD =2.81) (p <0.001), FlatDisharm - related (M = -0.28 mV, SD = 2.86) (p <0.001), and Real Word - related (M = -0.16 mV, SD = 2.23) (p <0.001). FlatHarm-unrelated (M = -0.67 mV, SD = 2.46) also had a more negative mean N400 amplitude than FlatHarm-related (p = 0.049), FlatDisharm-related (p = 0.013), and Real Word - related (p <0.001). However, the N400 amplitude for FlatDisharm-unrelated (M = 0.62 mV, SD = 2.88) reached significance only in comparison to Real Word - related (p = 0.002). Therefore, only FlatDisharm-unrelated trials failed to lead to a significantly lower N400 amplitude than all other related trials.

There was a significant effect of anteriority (F(1.138, 52.355) = 4.28,p = .04, η^2p = 0.085), where central regions (M = -0.94 mV, sd = 2.01) had a more negative amplitude than frontal regions (M = 0.23 mV, sd = 3.05). There were no significant differences between posterior regions (M = -0.73 mV, sd = 2.5) and

Table 3.3: Mean N400 amplitudes for each Flatness Harmony-Category Congruency condition and Minimum and Maximum mean amplitudes.

Conditions	Mean (sd)	Min / Max
FH-related	-0.32 (2.81)	-8.95 / 8.87
FH-unrelated	-0.67 (2.46)	-7.86 / 7.93
FD-related	-0.28 (2.86)	-8.34 / 10.27
FD-unrelated	-0.62 (2.88)	-10.77 / 11.67
Real Word - related	-0.16 (2.23)	-6.37 / 6.33
Real Word - unrelated	-0.83 (2.28)	-7.43 / 6.69

Note. All values are in mV.
Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

frontal regions or posterior regions and central regions.

The main effect of lateralization was also present ($F(1.833, 84.312) = 28.47, p < .001$, η^2p = 0.382). The mean amplitude for the midline region (M = -1.148 mV, sd = 2.78) was more negative than the left (M = -0.42 mV, sd = 2.26)($p < .001$) and right side (M = 0.133 mV, sd = 2.58)($p < .001$). The left side was significantly more negative than the right side ($p = .005$).

The effect of the Backness Harmony factor was significant ($F(1,46) = 6.018, p = .018$, η^2p = 0.116). Backness Disharmonious condition (M = -0.768 mV, SD = 2.749) led to a more negative-going brain activity pattern compared to the Backness Harmonious condition (M = -0.161 mV, SD = 2.408)($p = .018$).

In addition to the main effects, the ANOVA results indicated an interaction of anteriority and Flatness Harmony-Category Congruency ($F(4.682, 215.365) = 4.542, p < .001$, η^2p = 0.09).

The interaction was driven by more negative N400 mean amplitude for Real Word - unrelated conditions than FlatHarm-related and FlatDisharm-related conditions in the frontal regions. There were also more negative N400 mean amplitudes in the novel word conditions (FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated) compared to Real Word - related conditions in the posterior regions. See Table 3.4 for the mean and standard deviation values of condition and anteriority interaction; and Table 3.5 for the full list of significant differences.

Table 3.4: Mean N400 amplitudes for each Flatness Harmony-Category Congruency and anteriority interaction and Minimum and Maximum mean amplitudes

Flatness Harmony-Category Congruency	Anteriority	Mean (SE)	Min / Max
FH-related	F	0.82 (3.18)	-5.69 / 8.87
	C	-0.76 (2.13)	-5.73 / 4.58
	P	-1.02 (2.69)	-8.95 / 6.22
FH-unrelated	F	0.00 (2.92)	-7.86 / 7.93
	C	-1.06 (1.95)	-5.94 / 3.76
	P	-0.96 (2.29)	-6.66 / 4.17
FD-related	F	0.85 (3.29)	-7.60 / 10.27
	C	-0.72 (2.24)	-8.34 / 5.25
	P	-0.95 (2.62)	-7.22 / 4.75
FD-unrelated	F	0.26 (3.56)	-10.77 / 11.67
	C	-1.14 (2.06)	-7.87 / 4.89
	P	-0.96 (2.63)	-6.68 / 5.90
Real Word-related	F	-0.03 (2.57)	-4.93 / 6.33
	C	-0.53 (1.76)	-6.37 / 4.55
	P	0.09 (2.24)	-4.86 / 4.71
Real Word-unrelated	F	-0.49 (2.48)	-7.43 / 6.69
	C	-1.41 (1.82)	-7.38 / 2.26
	P	-0.58 (2.38)	-6.34 / 5.89

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. All values are in mV.

Table 3.5: Post Hoc Comparisons - anteriority x Flatness Harmony-Category Congruency

		Mean Difference	SE	t	pbonf	
F, FH - related	P, FH - related	1.849	0.5	3.698	0.045	*
	C, FH - unrelated	1.872	0.488	3.839	0.028	*
	C, FD - unrelated	1.945	0.488	3.988	0.016	*
	P, FD - unrelated	1.805	0.488	3.701	0.046	*
	F, Real Word - unrelated	1.302	0.259	5.03	<.001	***
	C, Real Word - unrelated	2.224	0.488	4.56	0.002	**
P, FH - related	F, FD - related	-1.883	0.488	-3.861	0.026	*
	P, Real Word - related	-1.121	0.259	-4.333	0.003	**
C, FH - unrelated	F, FD - related	-1.907	0.488	-3.909	0.021	*
P, FH - unrelated	F, FD - related	-1.825	0.488	-3.742	0.039	*
	P, Real Word - related	-1.064	0.259	-4.11	0.007	**
F, FD - related	C, FD - unrelated	1.979	0.488	4.058	0.012	*
	P, FD - unrelated	1.839	0.488	3.771	0.036	*
	F, Real Word - unrelated	1.336	0.259	5.163	<.001	***
	C, Real Word - unrelated	2.259	0.488	4.631	0.001	**
P, FD - related	P, Real Word - related	-1.05	0.259	-4.059	0.008	**
	P, Real Word - related	-1.077	0.259	-4.164	0.005	**

* p <.05, ** p <.01, *** p <.001

Note. P-value adjusted for comparing a family of 153

Note. Results are averaged over the levels of: Conditions, lateralization

The interaction between lateralization and Flatness Harmony-Category Congruency was also significant ($F(6.865,\ 315.799) = 2.54, p = .016,\ \eta^2\text{p} = 0.052$). Inspecting the results, the interaction was caused by the differences between conditions in the midline section. Although the midline section had the most negative N400 amplitude in all conditions, the mean amplitudes of the midline electrodes were different based on the Flatness Harmony-Category Congruency factor. See Table 3.6 for the mean and standard deviation values of condition and Table 3.7 lateralization interaction.

Table 3.6: Mean N400 amplitudes for each Flatness Harmony-Category Congruency and lateralization interaction and Minimum and Maximum mean amplitudes

Flatness Harmony-Category Congruency	Lateralization	Mean (SE)	Min / Max
FH-related	L	-0.29 (2.39)	-6.07 / 6.00
	M	-1.10 (2.98)	-8.95 / 6.64
	R	0.44 (2.84)	-7.22 / 8.87
FH-unrelated	L	-0.60 (2.09)	-5.22 / 5.12
	M	-1.41 (2.65)	-7.86 / 7.93
	R	0.00 (2.42)	-7.37 / 6.81
FD-related	L	-0.26 (2.48)	-7.22 / 7.17
	M	-0.81 (3.12)	-8.34 / 10.27
	R	0.25 (2.87)	-7.60 / 8.80
FD-unrelated	L	-0.52 (2.56)	-7.37 / 7.26
	M	-1.37 (3.04)	-10.58 / 11.67
	R	0.04 (2.86)	-10.77 / 8.91
Real Word-related	L	-0.23 (2.02)	-4.69 / 4.71
	M	-0.60 (2.35)	-6.37 / 6.28
	R	0.37 (2.21)	-4.81 / 6.33
Real Word-unrelated	L	-0.59 (1.94)	-5.94 / 5.51
	M	-1.59 (2.48)	-7.43 / 6.69
	R	-0.30 (2.19)	-5.97 / 5.19

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. All values are in mV.

Table 3.7: Post Hoc Comparisons - lateralization x Flatness Harmony-Category Congruency

		Mean Difference	SE	t	pbonf	
R, FH - related	M, FH - related	0.808	0.205	3.933	0.018	*
	M, FH - unrelated	1.122	0.226	4.971	<.001	***
	M, FD - unrelated	1.064	0.226	4.716	<.001	***
	M, Real Word - unrelated	1.293	0.226	5.728	<.001	***
M, FH - related	L, FH - related	-1.53	0.205	-7.451	<.001	***
	L, FH - unrelated	-1.091	0.226	-4.833	<.001	***
	R, FD - related	-0.841	0.226	-3.727	0.036	*
	L, FD - related	-1.339	0.226	-5.931	<.001	***
	L, FD - unrelated	-1.144	0.226	-5.068	<.001	***
	R, Real Word - related	-0.869	0.226	-3.851	0.023	*
	L, Real Word - related	-1.459	0.226	-6.463	<.001	***
L, FH - related	R, FH - unrelated	1.035	0.226	4.584	0.001	**
	M, FH - unrelated	1.844	0.226	8.172	<.001	***
	M, FD - related	1.228	0.226	5.439	<.001	***
	R, FD - unrelated	0.951	0.226	4.216	0.005	**
	M, FD - unrelated	1.787	0.226	7.917	<.001	***
	M, Real Word - related	1.019	0.226	4.515	0.001	**
	R, Real Word - unrelated	1.025	0.226	4.542	0.001	**
	M, Real Word - unrelated	2.015	0.226	8.929	<.001	***
	L, Real Word - unrelated	0.732	0.159	4.603	<.001	***
R, FH - unrelated	M, FH - unrelated	0.81	0.205	3.943	0.017	*
	L, FD - related	-0.843	0.226	-3.735	0.035	*
	L, Real Word - related	-0.963	0.226	-4.267	0.004	**
	M, Real Word - unrelated	0.981	0.226	4.345	0.003	**
M, FH - unrelated	L, FH - unrelated	-1.405	0.205	-6.842	<.001	***
	R, FD - related	-1.155	0.226	-5.119	<.001	***
	M, FD - related	-0.617	0.159	-3.877	0.018	*
	L, FD - related	-1.653	0.226	-7.324	<.001	***
	R, FD - unrelated	-0.893	0.226	-3.957	0.015	*
	L, FD - unrelated	-1.458	0.226	-6.461	<.001	***
	R, Real Word - related	-1.183	0.226	-5.244	<.001	***
	M, Real Word - related	-0.825	0.159	-5.188	<.001	***
	L, Real Word - related	-1.773	0.226	-7.856	<.001	***
	L, Real Word - unrelated	-1.112	0.226	-4.927	<.001	***
L, FH - unrelated	M, FD - unrelated	1.347	0.226	5.97	<.001	***
	M, Real Word - unrelated	1.576	0.226	6.982	<.001	***
R, FD - related	M, FD - unrelated	1.098	0.226	4.864	<.001	***
	M, Real Word - unrelated	1.326	0.226	5.876	<.001	***
M, FD - related	L, FD - related	-1.036	0.205	-5.045	<.001	***
	L, FD - unrelated	-0.841	0.226	-3.728	0.036	*
	L, Real Word - related	-1.156	0.226	-5.123	<.001	***
	M, Real Word - unrelated	0.788	0.159	4.951	<.001	***
L, FD - related	M, FD - unrelated	1.595	0.226	7.069	<.001	***
	M, Real Word - related	0.827	0.226	3.666	0.045	*
	R, Real Word - unrelated	0.834	0.226	3.694	0.041	*
	M, Real Word - unrelated	1.824	0.226	8.081	<.001	***

R, FD - unrelated	M, FD - unrelated	0.835	0.205	4.068	0.011	*
	L, Real Word - related	-0.88	0.226	-3.899	0.019	*
	M, Real Word - unrelated	1.064	0.226	4.713	<.001	***
M, FD - unrelated	L, FD - unrelated	-1.401	0.205	-6.82	<.001	***
	R, Real Word - related	-1.126	0.226	-4.989	<.001	***
	M, Real Word - related	-0.768	0.159	-4.826	<.001	***
	L, Real Word - related	-1.715	0.226	-7.601	<.001	***
	L, Real Word - unrelated	-1.055	0.226	-4.672	<.001	***
L, FD - unrelated	M, Real Word - unrelated	1.629	0.226	7.218	<.001	***
R, Real Word - related	M, Real Word - unrelated	1.354	0.226	6	<.001	***
M, Real Word - related	L, Real Word - related	-0.948	0.205	-4.614	0.001	**
	M, Real Word - unrelated	0.996	0.159	6.262	<.001	***
L, Real Word - related	R, Real Word - unrelated	0.954	0.226	4.226	0.005	**
	M, Real Word - unrelated	1.944	0.226	8.612	<.001	***
	L, Real Word - unrelated	0.661	0.159	4.154	0.006	**
R, Real Word - unrelated	M, Real Word - unrelated	0.99	0.205	4.821	<.001	***
M, Real Word - unrelated	L, Real Word - unrelated	-1.283	0.205	-6.247	<.001	***

* p <.05, ** p <.01, *** p <.001

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. P-value adjusted for comparing a family of 153

Note. Results are averaged over the levels of: Conditions, anteriority

There was no three-way interaction or interaction involving the Backness Harmony factor.

Overall, the N400 effect appeared for the Real Word - unrelated and FlatHarm - unrelated trials compared to all other related conditions. However, the FlatDisharm-unrelated condition failed to lead to a significant N400 negativity compared FlatHarm -related and FlatDisharm - related conditions, the only difference appeared for the Real Word - related condition. Moreover, there is a difference of the localization of the N400 ERP for the novel words and the real words. While the Real Word - related and Real Word - unrelated conditions had consistent N400 amplitudes throughout all anteriority conditions, the N400 signal was the most prominent in the frontal regions for the novel words.

LPC

Figure 3.5 shows the ERP waveform for Cz electrode with marked LPC time window.. A repeated measures ANOVA with conditions (6 levels: FlatHarm-related, FlatHarm-unrelated, FlatDisharm-related, FlatDisharm-unrelated, Real

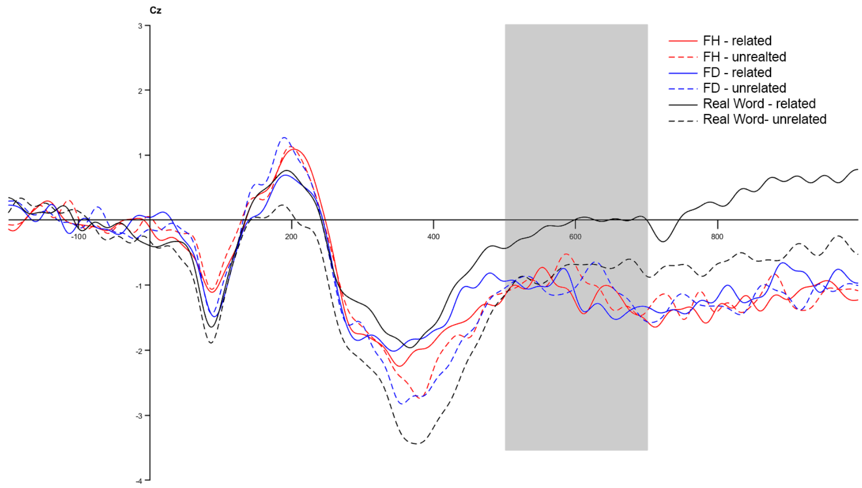

Figure 3.5: Averaged ERP waveforms for Cz electrode. LPC time window is marked

Word - related, and Real Word - unrelated), anteriority (3 levels: frontal, central, posterior), and lateralization (3 levels: left, midline, right) as within-subject factors and Backness harmony (2 levels: Backness Harmonious, Backness Disharmonious) as a between-subject factor was conducted for LPC measurements.

According to the ANOVA results, there was a significant main effect of conditions ($F(3.922,\ 180.401) = 5.85, p < .001,\ \eta^2\text{p} = 0.113$).

Follow-up pairwise tests revealed that the Real Word - related (M= 0.04 mV, sd = 2.05) had a significantly positive amplitude than novel word conditions (FlatHarm-related (M=-0.34 mV, sd=2.71) (p = .035), FlatHarm-unrelated (M = -0.58 mV, sd = 2.45) (p <.001), FlatDisharm-related (M = -0.42 mV, sd = 2.51) (p = .004), FlatDisharm-unrelated (M = -0.38 mV, sd = 2.62) (p = .011)). There was also a significant difference between Real Word - unrelated (M = -0.21 mV, sd = 2.08) condition and FlatHarm-unrelated(p = 0.05) condition where Real Word - unrelated had a more positive amplitude than FlatHarm - unrelated.

There were no main effects of anteriority, but the main effect of lateralization was significant (($F(1.836,\ 84.44\ ) = 34.032, p < .001,\ \eta^2\text{p} = 0.425$)). The right side (M = 0.48 mV, sd = 2.37) had significantly higher amplitudes compared to the left side (M = -0.78 mV, sd = 2.18) (p <.001) and the midline (M = -0.64 mV, sd = 2.51) (p <.001). The difference between the left side and midline was not significant.

There was a marginally significant effect of the Backness Harmony factor on the LPC signal ($F(1, 46) = 4.02, p = .051$, η^2p = 0.08). The Backness Harmonious condition (M = -0.0881 mV, sd = 2.279) had more positive amplitude compared to the Backness Disharmonious condition (M = -0.524 mV, sd = 2.528).

In addition to the main effects, the ANOVA results revealed an interaction between conditions and anteriority ($F(4.909, 225.832) = 3.171, p = 0.009$, η^2p = 0.064).

According to the follow-up post hoc comparisons, this interaction was driven by the more positive amplitude of Real Word - related condition in posterior regions compared to FlatHarm-related (p =.003), FlatHarm-unrelated(p = .029) and FlatDisharm-unrelated (p = .002) conditions. See Table 3.8 for the mean and standard deviation of the condition and anteriority interaction and Table 3.9 for the full list of significant differences.

Table 3.8: Mean LPC amplitudes for each Flatness Harmony-Category Congruency anteriority interaction and Minimum and Maximum mean amplitudes

Flatness Harmony-Category Congruency	Anteriority	Mean (SE)	Min / Max
FH-related	F	0.19 (3.17)	-7.97 / 9.59
	C	-0.51 (2.16)	-6.50 / 3.93
	P	-0.70 (2.65)	-6.77 / 5.45
FH-unrelated	F	-0.60 (3.03)	-8.50 / 7.64
	C	-0.60 (2.04)	-6.11 / 5.10
	P	-0.54 (2.16)	-6.02 / 4.70
FD-related	F	-0.39 (2.97)	-9.07 / 7.41
	C	-0.64 (2.12)	-6.19 / 4.68
	P	-0.22 (2.36)	-6.15 / 5.42
FD-unrelated	F	0.12 (3.30)	-9.69 / 10.17
	C	-0.57 (1.84)	-4.85 / 7.04
	P	-0.70 (2.47)	-5.83 / 7.00
Real Word-related	F	-0.43 (2.29)	-5.47 / 6.98
	C	0.12 (1.66)	-3.66 / 4.54
	P	0.44 (2.06)	-5.19 / 4.72
Real Word-unrelated	F	-0.35 (2.34)	-5.81 / 5.31

	C	-0.28 (1.78)	-5.63 / 5.16
	P	-0.01 (2.08)	-4.66 / 5.20

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. All values are in mV.

Table 3.9: Post Hoc Comparisons - anteriority x Flatness Harmony-Category Congruency

		Mean Difference	SE	t	Cohen's d	pbonf	
P, FH - related	P, Real Word - related	-1.152	0.266	-4.328	-0.495	0.003	**
P, FH - unrelated	P, Real Word - related	-1	0.266	-3.757	-0.43	0.029	*
P, FD - unrelated	P, Real Word - related	-1.158	0.266	-4.351	-0.498	0.002	**

* p <.05, ** p <.01, *** p <.001

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. P-value adjusted for comparing a family of 153

Note. Results are averaged over the levels of: Conditions, lateralization

The ANOVA results also showed interaction between conditions and lateralization ($F(7.102,\ 326.686) = 2.188, p = 0.034$, η^2p $= 0.045$). Post hoc comparisons revealed a significant pattern following the main effect of lateralization. Accordingly, the right side was generally significantly more positive than the midline and the left sides. However, The Real Word - related condition was an exception to this as it had a higher value in the midline compared to some other conditions, which resulted in non-significant comparisons. See Table 3.10 for the mean and standard deviation values of condition and lateralization interaction; and Table 3.11 for the full list of significant differences.

Table 3.10: Mean LPC amplitudes for each Flatness Harmony-Category Congruency and lateralization interaction and Minimum and Maximum mean amplitudes

Flatness Harmony-Category Congruency	Lateralization	Mean (SE)	Min / Max
FH-related	L	-0.86 (2.36)	-7.17 / 4.48
	M	-0.85 (2.77)	-7.97 / 7.00
	R	0.69 (2.71)	-6.97 / 9.59
FH-unrelated	L	-0.97 (2.26)	-7.68 / 4.70
	M	-0.97 (2.60)	-8.50 / 7.64
	R	0.20 (2.29)	-7.40 / 6.81
FD-related	L	-0.91 (2.23)	-9.07 / 5.42
	M	-0.76 (2.62)	-8.24 / 7.00
	R	0.41 (2.46)	-7.69 / 7.41
FD-unrelated	L	-0.90 (2.38)	-8.66 / 7.00
	M	-0.73 (2.62)	-9.69 / 10.17
	R	0.49 (2.66)	-9.04 / 7.71

Real Word-related	L	-0.46 (1.89)	-4.51 / 4.72
	M	-0.06 (2.09)	-5.47 / 4.66
	R	0.65 (2.01)	-4.53 / 6.98
Real Word-unrelated	L	-0.60 (1.88)	-5.81 / 3.85
	M	-0.48 (2.21)	-5.63 / 4.92
	R	0.45 (2.00)	-4.37 / 5.31

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. All values are in mV.

Table 3.11: Post Hoc Comparisons - lateralization x Flatness Harmony-Category Congruency

		Mean Difference	SE	t	Cohen's d	pbonf	
L, FH - related	R, FH - related	-1.547	0.214	-7.239	-0.664	<.001	***
	R, FH - unrelated	-1.056	0.232	-4.545	-0.454	0.001	**
	R, FD - related	-1.267	0.232	-5.449	-0.544	<.001	***
	R, FD - unrelated	-1.345	0.232	-5.788	-0.578	<.001	***
	R, Real Word - related	-1.496	0.232	-6.437	-0.643	<.001	***
	R, Real Word - unrelated	-1.301	0.232	-5.597	-0.559	<.001	***
M, FH - related	R, FH - related	-1.533	0.214	-7.173	-0.658	<.001	***
	R, FH - unrelated	-1.042	0.232	-4.485	-0.448	0.002	**
	R, FD - related	-1.253	0.232	-5.389	-0.538	<.001	***
	R, FD - unrelated	-1.331	0.232	-5.728	-0.572	<.001	***
	M, Real Word - related	-0.792	0.172	-4.6	-0.34	<.001	***
	R, Real Word - related	-1.482	0.232	-6.377	-0.637	<.001	***
	R, Real Word - unrelated	-1.287	0.232	-5.537	-0.553	<.001	***
R, FH - related	L, FH - unrelated	1.66	0.232	7.141	0.713	<.001	***
	M, FH - unrelated	1.66	0.232	7.14	0.713	<.001	***
	L, FD - related	1.598	0.232	6.874	0.686	<.001	***
	M, FD - related	1.44	0.232	6.195	0.619	<.001	***
	L, FD - unrelated	1.591	0.232	6.845	0.683	<.001	***
	M, FD - unrelated	1.418	0.232	6.098	0.609	<.001	***
	L, Real Word - related	1.141	0.232	4.907	0.49	<.001	***
	L, Real Word - unrelated	1.29	0.232	5.551	0.554	<.001	***
	M, Real Word - unrelated	1.166	0.232	5.018	0.501	<.001	***
L, FH - unrelated	R, FH - unrelated	-1.17	0.214	-5.473	-0.502	<.001	***
	R, FD - related	-1.38	0.232	-5.936	-0.593	<.001	***
	R, FD - unrelated	-1.459	0.232	-6.274	-0.626	<.001	***
	M, Real Word - related	-0.919	0.232	-3.955	-0.395	0.015	*
	R, Real Word - related	-1.609	0.232	-6.924	-0.691	<.001	***
	R, Real Word - unrelated	-1.414	0.232	-6.084	-0.607	<.001	***
M, FH - unrelated	R, FH - unrelated	-1.169	0.214	-5.473	-0.502	<.001	***
	R, FD - related	-1.38	0.232	-5.935	-0.593	<.001	***
	R, FD - unrelated	-1.458	0.232	-6.274	-0.626	<.001	***
	M, Real Word - related	-0.919	0.172	-5.337	-0.395	<.001	***
	R, Real Word - related	-1.609	0.232	-6.923	-0.691	<.001	***
	R, Real Word - unrelated	-1.414	0.232	-6.083	-0.607	<.001	***
R, FH - unrelated	L, FD - related	1.108	0.232	4.765	0.476	<.001	***
	M, FD - related	0.95	0.232	4.085	0.408	0.009	**

	L, FD - unrelated	1.101	0.232	4.736	0.473	<.001	***
	M, FD - unrelated	0.927	0.232	3.989	0.398	0.013	*
L, FD - related	R, FD - related	-1.318	0.214	-6.167	-0.566	<.001	***
	R, FD - unrelated	-1.396	0.232	-6.008	-0.6	<.001	***
	M, Real Word - related	-0.857	0.232	-3.688	-0.368	0.041	*
	R, Real Word - related	-1.547	0.232	-6.657	-0.665	<.001	***
	R, Real Word - unrelated	-1.352	0.232	-5.817	-0.581	<.001	***
M, FD - related	R, FD - related	-1.16	0.214	-5.428	-0.498	<.001	***
	R, FD - unrelated	-1.239	0.232	-5.328	-0.532	<.001	***
	M, Real Word - related	-0.699	0.172	-4.061	-0.3	0.008	**
	R, Real Word - related	-1.389	0.232	-5.977	-0.597	<.001	***
	R, Real Word - unrelated	-1.194	0.232	-5.137	-0.513	<.001	***
R, FD - related	L, FD - unrelated	1.311	0.232	5.64	0.563	<.001	***
	M, FD - unrelated	1.137	0.232	4.893	0.489	<.001	***
	L, Real Word - related	0.861	0.232	3.702	0.37	0.039	*
	L, Real Word - unrelated	1.01	0.232	4.346	0.434	0.003	**
	M, Real Word - unrelated	0.886	0.232	3.813	0.381	0.025	*
L, FD - unrelated	R, FD - unrelated	-1.39	0.214	-6.504	-0.597	<.001	***
	M, Real Word - related	-0.851	0.232	-3.659	-0.365	0.046	*
	R, Real Word - related	-1.541	0.232	-6.628	-0.662	<.001	***
	R, Real Word - unrelated	-1.345	0.232	-5.788	-0.578	<.001	***
M, FD - unrelated	R, FD - unrelated	-1.216	0.214	-5.691	-0.522	<.001	***
	M, Real Word - related	-0.677	0.172	-3.931	-0.291	0.014	*
	R, Real Word - related	-1.367	0.232	-5.881	-0.587	<.001	***
	R, Real Word - unrelated	-1.172	0.232	-5.041	-0.503	<.001	***
R, FD - unrelated	L, Real Word - related	0.939	0.232	4.041	0.403	0.01	*
	L, Real Word - unrelated	1.089	0.232	4.684	0.468	<.001	***
	M, Real Word - unrelated	0.965	0.232	4.152	0.415	0.007	**
L, Real Word - related	R, Real Word - related	-1.09	0.214	-5.102	-0.468	<.001	***
	R, Real Word - unrelated	-0.895	0.232	-3.85	-0.384	0.022	*
R, Real Word - related	L, Real Word - unrelated	1.24	0.232	5.334	0.533	<.001	***
	M, Real Word - unrelated	1.116	0.232	4.801	0.479	<.001	***
L, Real Word - unrelated	R, Real Word - unrelated	-1.045	0.214	-4.888	-0.449	<.001	***
M, Real Word - unrelated	R, Real Word - unrelated	-0.921	0.214	-4.309	-0.395	0.004	**

* p <.05, ** p <.01, *** p <.001

Abbreviations. FH, Flatness Harmonious; FD, Flatness Disharmonious

Note. P-value adjusted for comparing a family of 153

Note. Results are averaged over the levels of: Conditions, anteriority

There was no three-way interaction and no interaction with the between-participant Backness Harmony factor.

Overall, the LPC results showed a difference between Real Word - related conditions and the novel word conditions (FlatHarm-related FlatHarm-unrelated, FlatDisharm-related, and FlatDisharm-unrelated). The Real Word - related condition resulted in a more positive mean amplitude compared to novel words in the

LPC time slot. When the interaction between anteriority and Flatness Harmony-Category Congruency was considered, the effect between Real Word - related and most novel word conditions was present in the posterior regions, except for not having any interaction with the FlatDisharm-related condition.

3.3 Discussion

Experiment 2 was conducted to investigate the possible differences in the N400 and LPC ERPs in response to vowel harmony conditions of newly learned novel words via CSWL. We expected ERP signals for vowel harmonious novel words to be more similar to Real words. On the other hand, novel words that do not follow the vowel harmony rules to be more difficult for native Turkish speakers, which would lead to differentiation from Real and vowel-harmonious novel words.

The results of the experiment implied a difference in participants' overt response and ERP signals depending on the vowel harmony conditions. The responses for both the recognition test and the SJT showed above-chance level results, which implies that participants learned word-referent pairs correctly. Even though there was no effect of vowel harmony on the recognition test accuracy, the results of the linear mixed model showed that participants were faster responding to the FlatDisharm word.

While the recognition test request participants to choose the correct referent picture, the SJT is a more cognitively demanding task that requires participants to recall the correct meaning of the words and compare the semantic category of it to the presented second word. Participants again showed an above-chance level accuracy, but more importantly, there were differences between vowel harmony conditions according to the models. The difference was usually due to higher accuracy FlatDisharm - unrelated condition while there were no such difference in the FlatHarm - unrelated condition.

For the ERP results, the main difference was in the N400 ERP time window.

The N400 results for the Real words were in line with the literature where Real Word - unrelated trials lead to more negative-going ERP responses compared to the Real Word - related trials. A similar N400 pattern was also seen between FlatHarm-unrelated and FlatHarm-related trials, where FlatHarm-unrelated is more negative going than FlatHarm-related. This implies that FlatHarm words were learned and processed similarly to real Turkish words. On the other hand, the same difference was not present between FlatDisharm-unrelated and FlatDisharm-related trials. However, FlatDisharm-unrelated trials were significantly more negative than Real Word - related trials.

The LPC results were led by the difference between Real Word - related conditions and the novel word conditions, where the Real Word-related condition led to a more positive going ERP signal. This difference is probably because Real Words were already known to the participants but they had very limited exposure to novel words.

The results of Experiment 2 will be discussed in more detail in the general discussion chapter.

Chapter 4

General Discussion

The main aim of this study was to investigate the effects of native language word characteristics on novel word learning and the associated neural mechanisms underlying novel word learning using a CSWL paradigm. In this study, we focused on how learning words that either do or do not obey Turkish vowel harmony rules affect native Turkish speakers' novel word learning. There are many studies on CSWL and vowel harmony separately investigating different manipulations of each, for example: manipulating the number of words and referents in CSWL paradigms (Mulak et al., 2019; Yurovsky and Frank, 2015), examining the similarity of words (such as minimal vs. non-minimal pairs) in CSWL paradigms (Escudero et al., 2016; Tuninetti et al., 2020), and, for vowel harmony work, focusing on the generalization of vowel harmony rules and identification of word boundaries (LaCross, 2015) and how speakers of non-harmonic languages learn and generalize vowel harmony rules (Finley, 2021; Martin and White, 2021). While most CSWL studies have examined cross-lexical differences with monosyllabic words, investigating vowel harmony necessarily implies within-lexical structures with multisyllabic words. This paper broadens existing knowledge in both areas by looking at both the effect of within-lexical properties on CSWL and the implication of native harmony patterns on learning.

To briefly recap, Turkish has two different vowel harmony rules. The Backness harmony rule, used as a between-subject factor in this design, is a simple harmony rule which requires all the vowels in a word to have the same back/front character. On the other hand, the Flatness harmony rule is more complex by changing the subsequent vowel according to both the roundness and height of the vowel. Previous studies have shown that when there is a simple vowel harmony rule like Backness harmony, even speakers of non-vowel harmonic languages recognized the pattern and generalized it to new situations, such as novel words in that language (Finley, 2012; Martin and Peperkamp, 2020; Skoruppa and Peperkamp, 2011). However, the complex vowel harmony rules, where the change in the vowel depends on multiple factors, were difficult for that same participant group (Skoruppa and Peperkamp, 2011). This difference, until now, had not been investigated in speakers of vowel harmonic languages. Although most of these studies were done with simple vowel harmony patterns, both Backness harmony and Flatness harmony were analyzed and controlled in this study.

Additionally, many of the previous studies had used only behavioral measurements, such as accuracy information or participant reaction time; recently, ERPs (measured with EEG) have been used to further probe the neural mechanisms underlying novel word learning and linguistic influences on this process (Angwin et al., 2022; Mangardich and Sabbagh, 2022). Here, two ERP responses associated with semantic consolidation, the N400 and the LPC, were the focus of this study to provide converging evidence of how participants' overt behavior is linked to underlying neural responses. The important difference between N400 and LPC responses were the elicit time period and possible cognitive differences in retrieval. Briefly, the N400 is related to the semantic congruency of the words and is associated with a more automatic response (Kutas and Federmeier, 2011) while the LPC is thought to rely on memory traces, suggesting that it is not as automatic (Rugg and Curran, 2007; Bakker et al., 2015).

The current project has three research questions; (1) Can participants learn both vowel harmonious and disharmonious words via CSWL? (2) Can novel words lead to N400 or LPC ERPS in case of semantic incongruency, and lastly? (3) would vowel harmony conditions of novel words have an effect on N400 and LPC ERPs. These questions were examined by manipulating novel bisyllabic words that conformed to Turkish phonotactics and varied systematically on their backness vs. flatness vowel harmony. The experiment consisted of three phases; training, recognition test, and SJT. Participants were presented with two novel words and two referent pictures in the training phase, then immediately started the recognition test. In the recognition test, participants were asked to indicate which of the referent pictures on the screen is the correct meaning of the presented word. Following that, the SJT was run where novel and real words were paired with real Turkish words from the same or different semantic categories, which resulted in six conditions (FlatHarm - related, FlatHarm - unrelated, FlatDisharm - related, FlatDisharm - unrelated, Real Word - related and Real Word - unrelated). Participants were asked to indicate if the words in the pairs were semantically related.

In regards to the first research question, considering the previous literature on CSWL and vowel harmony, which shows that this paradigm is very successful at having participants learn novel words, we expected participants to show above-chance performance in learning for all the tested words. However, since disharmonious words did not obey the features of the native language, we expected that participants would have difficulties in learning disharmonious words as characterized by lower accuracy in disharmonious words and significant reaction time differences between conditions. Layered into this base prediction, we also had different hypotheses for the Backness and Flatness harmony rules that were manipulated. Specifically, Backness harmony was a between-subject factor because it was not possible to include as a within-participant factor due to the constraints on building novel words that conformed to these two Turkish vowel harmony rules. Therefore every participant was exposed to a consistent Backness harmony rule, either harmonious or disharmonious; we did not predict a difference between conditions. This is not unlike previous studies that have shown that speakers of vowel harmonic languages tend to adapt to consistent patterns within the constraints of a learning scenario (Altan, 2011; LaCross, 2015); there was no reason to suspect there may be differences within participants with this factor.

On the other hand, since all the participants would be exposed to both

FlatHarm and FlatDisharm words (a within-participant factor), the Flatness harmony rule could not be generalized to the novel words as consistently. This would allow us to see the possible learning difficulty for the words contrasting with the native language rule, FlatDisharm words, when learners cannot adapt to a consistent vowel harmony rule. We believed that this learning difficulty would be seen in both the recognition test with a lower accuracy rate and in the Semantic judgment test by falsely identifying semantic categories for FlatDisharm words.

The other research question the current study tries to answer is whether novel words learned via CSWL lead to neural signals in response to semantic incongruences, such as the N400 and LPC ERP responses as markers of semantic consolidation. As recent research on CSWL shows an N400 robust when newly learned words were presented with words from unrelated semantic categories compared to when they were presented with words from a related semantic category suggesting a successful semantic consolidation (Angwin et al., 2022; Mangardich and Sabbagh, 2022), we expected to replicate and extend those results. Even though the research with CSWL did not measure the LPC signal, some other studies have shown that the LPC is another ERP signal that arises to semantic incongruency in the early phases of learning (Bakker et al., 2015; Lei et al., 2022; Liu and Hell, 2020). More importantly, those studies showed that the N400 and LPC ERPs are related to different phases of semantic consolidation, with the LPC being in the early learning phases where the associations are newly made and recalling is more cognitively demanding, whereas the N400 appears when the association between words and referent are stronger. Therefore, a difference in the LPC signal between related and unrelated pairs would mean the word requires explicit recalling, while if such a difference is present in the N400 phase, the connection between words and the referents starts to be automatized.

The third research question, the differences in the neural signals due to vowel harmony in newly learned words, was of particular interest. The hypothesis for this question was that the newly learned words that obey the Flatness harmony rule would have more similar neural signals to real words compared to disharmonious words. This hypothesis consists of multiple sub-hypotheses depending on the strength of semantic consolidation. If learners had a stronger semantic connection between words and referents, the difference should be seen at the N400 time window; however, if the connections were not strong enough and recalling the meaning requires controlled processing, the difference should be at the LPC time window. In both cases, we expect to see a robust N400 response for Real Word - unrelated condition compared to Real word - related condition as the Real word trials were the control group where the neural signals should be in line with the literature.

In the case of the first scenario, where the difference in the N400 response implies automatic recalling of the referents, we expected the largest N400 amplitude for real word-unrelated conditions and the smallest N400 amplitude for real word-related conditions as they are familiar words. We also expected to see a similar pattern to Real Word conditions for FlatHarm - related and FlatHarm - unrelated conditions where FlatHarm - unrelated condition has a significantly larger N400 mean amplitude than FlatHarm - related condition. According to our hypothesis, learners would have difficulty with learning FlatDisharm words due to not obeying their native language's rule; therefore, we expect that there would not be a significant difference between FlatDisharm - related and FlatDisharm - unrelated conditions at the N400 time window.

On the other hand, if the recalling of words' meanings is not automatic yet, a difference could appear in the LPC time window. The LPC signal arises in response to semantic congruency in the early learning phase, where the recalling of the meanings of the words is not automatic yet. Accordingly, there should not be an LPC ERP difference between the Real Word - related and Real Word-unrelated trials as they are well-known words that have automatic recalling. Since we expect FlatHarm words to have a stronger association with their referents due to following the rules of Turkish, we expected them to have a similar neural signal pattern to real words. Therefore, for the LPC time window, we did not expect a significant difference between FlatHarm - related and FlatHarm - unrelated conditions. However, we expected to see a difference between FlatDisharm - related and FlatDisharm - unrelated words where FlatDisharm - related condition have a larger mean LPC amplitude (more positive) compared to FlatDisharm -

unrelated condition.

In general, the results of the experiment generally confirmed our hypotheses. Participants showed above chance performance in the recognition test, implying participants were able to learn the words successfully. There was no difference between vowel harmony conditions in the recognition test accuracy, but there were differences in the SJT based on the vowel harmony conditions. Moreover, a higher amplitude N400 signal was seen for FlatHarm - unrelated compared to FlatHarm - related and RealWord- unrelated compared to Real word - related trials, no such difference was present for FlatDisharm - related and FlatDisharm - unrelated trials. There were no differences between harmony trials in the LPC signal, although all the novel word trials had significantly lower amplitude than Real Word-related trials. In the following section, we will discuss the implication of the results and how they fit into the current literature.

Behavioral Results

Recognition test

The results of the recognition test showed a ceiling effect on learning with a performance of around a 90% accuracy rate. This above-chance performance was expected and is in line with the current literature on CSWL. The results support the idea that even without explicit instructions, people can learn the word-referent pairs by tracking their co-occurrence in a CSWL paradigm.

The important feature of CSWL is mimicking naturalistic learning where the words were not directly one-to-one matched with their referents, as happens in explicit word learning. By not giving explicit instructions about the learning paradigm, simply by asking participants to watch the screen but not saying they should follow the words and referents, CSWL creates an environment more similar to how people would encounter novel words naturally. Previous studies showed that the ambiguity of the CSWL paradigm has an effect on learning, as the number of possible referents increases, the accuracy performance decreases (Mulak et al., 2019; Vouloumanos, 2008). We used the least ambiguous CSWL paradigm with 2 words and 2 referents in the training and recognition test to ensure that learning would take place.

Even though a high performance was expected, an accuracy rate of around 90% is higher than most previous CSWL studies, barring a few studies (Angwin et al., 2022). The major difference between the lower but still above-chance accuracy rate in the previous literature and our ceiling effect accuracy might be due to the type of referents used. In most of the previous literature, the referents were usually non-real objects which participants saw for the first time (e.g., nonce items) (Escudero et al., 2016; Mulak et al., 2019; Tuninetti et al., 2020). Because of the methodological requirements for the SJT here, the stimuli were real objects that participants were familiar with, such as four-legged animals or vehicles. This familiarity with the objects may have helped participants in learning because they did not need to process new objects as well as new words (see also Angwin et al., 2022 for similar results).

The other important aspect of the study was the effect of vowel harmony rules during learning as a within lexical unit manipulation. Most previous studies on CSWL have investigated the interaction of word boundaries with minimal/non-minimal word pairs in monosyllabic words. The literature suggests that minimal pairs are more difficult to distinguish compared to non-minimal pairs, as participants show a higher accuracy rate in the non-minimal pair conditions (Escudero et al., 2016; Tuninetti et al., 2020). However, when learning a language, the vocabulary does not only consist of monosyllabic words and may include within-lexical features present in multisyllabic words, such as vowel harmony. Here, we expected to see higher accuracy in the harmonious conditions as they are more similar to what the participants' native language has, and the participants can rely on existing L1 rules. However, the results of the recognition test did not support this hypothesis because it did not show any difference between vowel harmony conditions. The results of the recognition test showed very high accuracy, with 92% in the Backness Harmony condition and 89% in the Backness Disharmony condition. Although the performance was higher in the Backness harmony condition, the difference was not statistically significant. We also didn't see a difference between FlatHarm and Disharmonious stimuli in the recognition test accuracy.

It is likely that our non-significant result in the recognition test is due to a ceiling effect. As the participants were highly successful in learning the novel words in the experiment, the ceiling effect might have obscured the difference between harmonious and disharmonious words. Moreover, previous studies also showed that speakers of vowel harmonic languages are more prone to find word boundaries according to vowel harmony, regardless of whether the harmony pattern follows their native language's harmony pattern or not (Altan, 2011; LaCross, 2015). Here, backness harmony was a between-subject factor, such that participants heard the words either harmonious or disharmonious for the backness harmony. This may have led the participant groups to become habituated to either a consistent harmonious or consistent disharmonious backness pattern, and no difference was found between the two. Finally, although all participants were exposed to flatness harmony and disharmony as a within-participant factor, this also did not show significant differences in learning the two categories of words. While it is harder to speculate as to the reason why flatness did not show differences in learning performance, we suspect that participants were simply able to rely on existing L1 knowledge and existing familiarity with the referents in such a way that allowed them to accurately link the correct object-referent pairs, even when disharmonious.

Although there were no differences in the accuracy for vowel harmony conditions, a significant difference in reaction time was present in the recognition test. Participants were faster when identifying novel words that did not obey Flatness harmony. Although there is not much work on reaction time differences due to vowel harmony, a similar response pattern to disharmonious words was previously present in Perea et al. (2021). Different from our design, in their paper, a priming task was done with the harmonious and disharmonious orthographic words. After being primed with disharmonious words, participants were faster at identifying the words. Although the difference between the harmonious and disharmonious words was explained by the fact that the disharmonious words contained diacritics, our pseudowords were presented auditorily. However, there is literature linking orthography and phonology quite strongly, and so the shorter reaction times for disharmonious words might be a result of an unusual word structure compared to participants' Turkish L1.

Semantic Judgemewnt Task

The SJT was used to measure the level of semantic consolidation for the novel words; in this study, to be successful in the task, participants have to both access the semantic meanings of the novel words and be able to compare them to existing semantic categories. Overall, participants were successful in this task with above-chance response accuracy in all conditions.

As expected, the accuracy rate for the real words was higher than for the novel words in both related and unrelated trials. Moreover, there was a difference between novel words - related and novel words - unrelated trials, as participants were more successful in the unrelated trials. This accuracy pattern was seen in previous literature where participants showed higher accuracy for unrelated trials, suggesting a rejection bias (Dittinger et al., 2016; Angwin et al., 2022).

More interestingly, this rejection bias showed a difference between different vowel harmony trials. According to model 3 in accuracy, while participants had higher success in the FlatDisharm - unrelated trials than in both FlatHarm - related and FlatDisharm - related trials, such a difference was not present in the FlatHarm - unrelated trials. We interpret these results, seeing significantly higher accuracy on the FlatDisharm - unrelated than FlatDisharm - related trials, as participants having more difficulty accessing the meaning of the FlatDisharm word. To be successful in the SJT, the participant had to remember the correct referent for the word and compare it to the presented second word. For the unrelated trials, they do not have to remember the correct referent as long as they categorize those words in different semantic categories. However, for the related trials, recalling the correct referent is a must. This might be leading to bias towards rejection.

Finally, there was a clear effect of L1 phonological rules on the behavioral results. From the perspective of this study, these results imply that creating conceptual links between novel words and referents was easier when the words followed the rules of L1. This might be because participants were able to rely on the word structure of their L1 when forming connections between harmonious words and their meanings. However, the rules of the L1 did not help when presented with disharmonious words. This might have led to higher uncertainty in response to disharmonious words, which increased the rejection bias in disharmony trials. Such a difference was also supported by the ERP results, which will be discussed in more detail in the next section.

EEG measurements

The behavioral result of this study gave us valuable information about the participant's overt biases about the learning words with different vowel harmony rules; ERP measurements give a more sensitive view with respect to understanding word learning. Up to this date, very few studies examined the ERP responses after learning via CSWL (see also Angwin et al., 2022;Mangardich and Sabbagh, 2022). Similar to the previous ERP studies on word learning, we examined the mean amplitudes for the N400 and LPC responses in this study.

According to the ERP recordings, the main difference between vowel harmony trials appeared in the N400 time window rather than LPC, although there were some differences between novel words and real words in the LPC time window.

N400

The ERP results revealed an N400 increase for the unrelated trials in almost all of the conditions (FlatHarm, FlatDisharm, and Real Word). This robustness in the N400 is considered an effect of word learning in the previous literature, as the presence of the N400 difference shows that connections between words and referents were established. Although most of the previous research is on explicit learning paradigms in which words were matched with referents without any ambiguity, recent ERP studies using CSWL as an implicit learning paradigm also established an N400 effect (Angwin et al., 2022; Mangardich and Sabbagh, 2022).

Before discussing the influence of vowel harmony, the difference in the localization of the N400 effect gives us some insight for learning. In our study, the N400 potentials for the novel words were higher than the Real word trials in the frontal regions. Moreover, the amplitude difference between related trials and unrelated trials was significant only in this area for the novel words, while the N400 waveforms to the related vs unrelated trials were not significantly different from each other in the posterior region. On the other hand, real word trials did not show any difference in anterior regions; although the central electrodes had a higher N400 difference between real words related and unrelated trials, the N400 effect was present in all regions. This frontal activity pattern in response to newly learned words was also seen in the previous literature (Mestres-Misse et al., 2007). According to them, this suggests that the semantic connections for novel words were not as strong as the real words, which led to a higher cognitive control for the novel words with a higher activity of the frontal cortex. However, this frontal activity pattern was not present in all of the novel word learning studies. Angwin et al. (2022) did not observe any differences between brain regions.

The novel side of this study was the vowel harmony element of these newly learned words, and the results of the study supported our hypothesis that there would be an effect of vowel harmony on learning. For backness harmony, the Backnes Disharmony condition had a more negative N400 response compared to the Backness Harmony condition. It is difficult to speculate about the reason for this difference without separating the effect of relatedness; we believe it might be due to the backness disharmony pattern. As the N400 ERP is highly related to expectancy (Kutas and Federmeier, 2011), disharmonious words might be going against the expectancy of the participant for the word structure. Unfortunately, we cannot comment more on the relation between Backness harmony and relatedness with respect to N400 amplitude because Backness harmony was a between-subject factor, and the nature of the stimuli design based on the Turkish vowel harmony rules did not allow us to analyze that interaction.

Nonetheless, there was an effect of the within-subject conditions, Flatness Harmony and Category Congruency, on the N400 response. As expected, the control trials, Real word-related and Real word-unrelated, had the lowest and highest N400 amplitudes, respectively, among all the conditions. More importantly, the N400 amplitude of FlatHarm - unrelated trials was less negative than Real Word - unrelated trials but more negative compared to all other categories of related trials. However, FlatDisharm - unrelated trials did not lead to such a difference in the N400 response; it was only different from Real Word - related condition. This implies an effect of L1 rules on word learning because the N400 difference was more prevalent in the FlatHarm condition similar to the Real Word-related/unrelated difference. On the other hand, FlatDisharm trials did not reach significance for the N400 response, which might be due to added difficulty for participants since the stimuli were contrary to the rules of the native language. As it is supported by the behavioral results of the SJT, participants cannot rely on their L1 information for the disharmonious trials as happened for harmonious words, so participants have to be aware of other aspects of the words.

Comparing the N400 result for the Backness harmony and Flatness Harmony factors might be confusing, as while the Backness Disharmonious condition had more negative going signals, the FlatDisharm condition was more positive going compared to FlatHarm. However, it is important to note that these signals are not reflecting the same concept as the Flatness Harmony condition compares relatedness, while the Backness Harmony condition was the average of all conditions. Although comparing different type of factors may explain these differences, the differences might be due to the nature of vowel harmony type too. Since backness harmony is a simple vowel harmony rule, it it might be more salient, eliciting more attention from the participants, and therefore increasing the overall signal (leading to higher amplitude waveforms). On the other hand, flatness harmony is a complex vowel harmony rule which might not change the overall signal (as FlatHarm - related and FlatDisharm - related conditions were not significantly different than each other), but it might still affect learning as seen in the FlatDisharm - unrelated condition. Further research investigating the Backness Harmony separately as a within-subject condition could give better answers to these points.

LPC

Unlike the N400 results, the LPC results did not show a significant difference for the Flatness harmony rule. The only difference on the Flatness harmony - Congruency factor was a higher amplitude of Real Word - related trials over all the novel word trials independent of the relatedness. Moreover, there was also no difference between Real Word-related and Real Word-unrelated trials for the LPC response.

As in the N400 ERP, there was also a difference between Backness Harmonious and Backness Disharmonious conditions, with the Harmonious condition having a higher amplitude value. Similar to our comment for the N400 difference, Backness disharmonious words might have led to overall less positive signals because of going against the native language pattern. However, we should be cautious about this interpretation given the design of the experiment.

In the novel word learning literature, the LPC is known as a signal that elicits a higher amplitude when the presented words are semantically related (Bakker et al., 2015; Liu and Hell, 2020; Lei et al., 2022). The differentiation between the N400 and LPC seems to be influenced by the strength of the semantic connection between words and their meaning on this effect. In that sense, the LPC is assumed to be related to the strength of the semantic connection between newly learned words and their referents in early learning phases (Bakker et al., 2015). Accordingly, if the recall requires an explicit effort, the LPC signal would be higher for the presentation of two related words compared to unrelated words. However, as the recalling process gets easier, this effect in the LPC time window decreases, and the N400 effect for unrelated word pairs is elicited.

In this sense, our results could mean that the semantic connections between the words and the referents were strong enough to elicit an N400 response rather than LPC, as there were no differences between related and unrelated trials. Moreover, real words having a higher semantic strength between them might have led to higher potentials in the LPC effect. As the novel words were only learned with a couple of trials in the training section, the semantic connection for them was not established in high degrees compared to real words. With longer training distributed over multiple days, the difference between real words and novel words in the LPC time window might disappear.

Learning a foreign language requires a learner to adapt to the construction elements of the novel language from the perspective of their first language. In the sound perception aspect, there are different models, such as PAM and SLM models, claiming that sound categorization is highly affected by the learners' first language (Best and Tyler, 2007; Flege, 1995). Although from different perspectives, both models suggest that if the sound of the learned language is similar to the first language, categorization of the novel sound and, therefore, differentiation between the learned language's sounds would be easier. On the other hand, if there is no established similar sound in the first language, it would be more difficult to categorize that sound, which shows the effect of the first language on the learned language. Even though this study does not focus on individual sounds in a foreign language, it has a similar focus by suggesting there is an impact of the first language's rules on novel word learning. In our study, instead of focusing on individual sounds, we look at multi-syllable structures to see whether there is an effect of the first language's lexical constraints would have an effect on novel word learning.

Overall, the results show that L1 Turkish speakers use the Turkish phonological rules, vowel harmony, while learning novel words. Moreover, expanding the current literature, we see that both a simple rule like backness harmony and a more complex rule like flatness harmony showed a difference in both behavioral and ERP results. Previous studies show that speakers of non-harmonic languages can adapt to simple vowel harmony rules where there is a consistent pattern of a single feature (like height) (Finley, 2012; Skoruppa and Peperkamp, 2011) . However, French speakers, who do not have vowel harmony in their native language, had difficulty in generalizing that rule correctly when the rule changes according to multiple characteristics of a vowel in a complex vowel harmony rule, participants (Skoruppa and Peperkamp, 2011). Although this study did not focus on generalization of the rules, we saw an effect of flatness harmony, which can be considered a complex rule, on the novel word learning. Since Turkish speakers are exposed to both backness and flatness harmony rules in their native language, we believed they would show a difference between harmonious and disharmonious words, unlike what happened with French participants.

Such a difference for both backness harmony and flatness harmony indicates the effect of the native language. Previous studies with native speakers of a vowel-harmonic languages showed that they are better at finding word boundaries in a string of vowel-harmonious words (LaCross, 2015). With this study, we add that having such a rule in the native language also has an influence on the development of semantic connections between novel words and their referents. More importantly, we suggest that for speakers of Turkish even more complex harmony rules where there are multiple features in the vowel harmony rule can lead to such a difference. Although it is hard to comment on non significant differences, we believe that the lack of difference between the backness harmonious and the backness disharmonious words might be an extension of adapting and generalizing the vowel harmony rule. Since backness harmony is a simple rule, it might be easier for participants to generalize the rule for the language. On the other hand, we found a difference in flatness harmony, which was a within-subject factor, in both behavioral and ERP results. It is especially interesting considering flatness harmony is a complex rule depending on two characteristics of the vowel. For us, this shows a clear effect of native language's characteristics on learning novel words. Future research changing Backness and Flatness harmony factors might give more support to our interpretations.

One other thing we might need to consider to interpret the results is the language background of the participants. Although the ERP responses of the current study suggest a difference between related and unrelated trials for the N400 time window, the literature shows both N400 and LPC effects in different situations. For example, even though Bakker et al. (2015) found an N400 effect after a consolidation period, Lei et al. (2022) and Liu and Hell (2020) failed to observe the difference in the N400 while an LPC was present with the same experimental design. As it was discussed, the N400 is related to more automatic processing after stronger connections between words and referents. In that aspect, participants' language background might help to establish semantic connections between words and referents. In our experiment, similar to Bakker et al. (2015), participants knew at least two languages. In this study, they were proficient Turkish - English bilinguals who mostly started learning English in middle school. Therefore, this specific participant group has language learning experience, which might have helped them to learn semantic connections more efficiently. On the other hand, the result might have changed if the participant group was monolingual Turkish speakers. Whether an N400 robust would be seen in such a participant group would be an interesting topic for future research.

4.1 Limitations and Future Directions

As it is briefly discussed in the previous section, there are some limitations of this study affecting both first and second Experiments. For the First Experiment number of participants who were native speakers of non-harmonic languages was a limitation of the experiment, especially considering there was a variability in participants' language background. Even though we discarded the participant who had any knowledge of vowel harmonic languages, we could not control whether their native language had another effect on the results. Moreover, the number of participants was fewer than the Turkish participant group which created another limitation for the study. Doing the same experiment with a larger participant group of speakers of non-harmonic languages would improve the findings of the study.

In general, the participants' background can be considered a limitation of the study for Turkish participants for both Experiment 1 and Experiment 2. Since university students creates a highly educated participant group, we should be cautious to generalize the result of the study. Moreover, as it was mentioned, this participant group is highly proficient in their L2 (English). Their language learning experience might have helped them in the experiments. Having a different participant group who are monolingual Turkish speakers should be examined in future studies. We believe especially the ERP results might have influenced by the participant group. Considering the N400 ERP suggest a more automatic response, participants' previous language learning experience might have helped them; yet, in a monolingual group the LPC ERP might be present instead of the

N400.

Lastly, the number of novel words that were trained in the experiments were very limited with eight words in between-subject each condition. Although the design of the experiment was adapted from (Angwin et al., 2022) and the number of stimuli was the same as previous CSWL studies (Escudero et al., 2016; Mulak et al., 2019; Tuninetti et al., 2020), the number of words per vowel harmony condition was half of it with four words and each word had a different consonant pair. As the recognition test results from the both Experiment 1 and Experiment 2 showed very high accuracy performance, our task was easy for the participants. Since our study suggests that there is an effect of vowel harmony on word learning via CSWL with this limited number of words, doing a similar experiment by including more words and using the same consonant pairs for both harmonious and disharmonious words in the same condition would be a novel and interesting future study.

4.2 Conclusion

Overall, this study investigated the effect of Turkish vowel harmony on word learning via CSWL and whether the native language's phonetic rules would lead to differences for native Turkish participants. Although we could not see an effect of vowel harmony on the recognition test accuracy, the results revealed an effect of this factor in the SJT. Supporting our hypothesis, the result for the harmonious words followed a similar pattern to the real words in the ERP results. This implies that both real words and harmonious novel words were processed similarly after a brief learning period.

The results of the study shows that there is an effect of L1 within lexical unit items on multi-syllabic word learning. Although this study investigates this effect after a short training, we believe the results would be similar in other second language learning situations, especially in more naturalistic learning situations where people learn novel words by encountering it multiple times in different situations (e.g. hearing it on TV shows instead of studying a word list) similar to CSWL paradigm. Therefore, from a more general perspective, this study shows that people would have stronger semantic connections for the words that obey their native languages' within-lexical unit characteristics earlier compared to the words that do not follow those characteristics. In other words, people can learn words that have similar lexical patterns to their native language than the dissimilar ones when learning in natural situations.

References

Altan, A. (2011). Yapay bir dil ediniminde Ünlü uyumu kuralının türkçeyi ana dil olarak konuşanlardaki rolü. *Hacettepe Üniversitesi Türkiyat Araştırmaları (HÜTAD)*, 14(14):27–43.

Angwin, A. J., Armstrong, S. R., Fisher, C., and Escudero, P. (2022). Acquisition of novel word meaning via cross situational word learning: An event-related potential study. *Brain and Language*, 229:105111.

Angwin, A. J., Phua, B., and Copland, D. A. (2014). Using semantics to enhance new word learning: An erp investigation. *Neuropsychologia*, 59:169–178.

Audacity Team (1999-2021). Audacity® software.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., and McQueen, J. M. (2015). Tracking lexical consolidation with erps: Lexical and semantic-priming effects on n400 and lpc responses to newly-learned words. *Neuropsychologia*, 79:33–41.

Battig, W. F. and Montague, W. E. (1969). Category norms of verbal items in 56 categories a replication and extension of the connecticut category norms. *Journal of experimental Psychology*, 80(3p2):1.

Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., and Domínguez, A. (2018). Brain signatures of new (pseudo-) words: Visual repetition in associative and non-associative contexts. *Frontiers in Human Neuroscience*, 12:354.

Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., and Domínguez, A. (2019). Novel word learning: Event-related brain potentials reflect pure lexical and task-related effects. *Frontiers in Human Neuroscience*, 13:347.

Bermúdez-Margaretto, B., Shtyrov, Y., Beltrán, D., Cuetos, F., and Domínguez, A. (2020). Rapid acquisition of novel written word-forms: Erp evidence. *Behavioral and Brain Functions*, 16(1):11.

Best, C. T. and Tyler, M. D. (2007). *Nonnative and second-language speech perception: Commonalities and complementarities*, volume 17, page 13–34. John Benjamins Publishing Company, Amsterdam.

Boersma, P. and van Heuven, V. (2001). Speak and unspeak with praat. 5(9).

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4):433–436.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., and Lepage, M. (2010). The bank of standardized stimuli (boss), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE*, 5(5):e10773.

Cinan, S. (2001). 56 kategorİden kelİmelerİn frekanslari. *Psikoloji Çalışmaları*, 22:155–226.

Darcy, I., Mora, J. C., and Daidone, D. (2016). The role of inhibitory control in second language phonological processing: Inhibitory control and l2 phonology. *Language Learning*, 66(4):741–773.

Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., and Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, 28(10):1584–1602.

Do, Y. and Mooney, S. (2022). Variation awaiting bias: Substantively biased learning of vowel harmony variation. *Journal of Child Language*, 49(2):397–407.

Escudero, P., Mulak, K. E., and Vlach, H. A. (2016). Cross-situational learning of minimal word pairs. *Cognitive Science*, 40(2):455–465.

Finley, S. (2012). Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes*, 27(10):1550–1562.

Finley, S. (2021). Learning exceptions in phonological alternations. *Language and Speech*, 64(4):991–1017.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92:233–277.

Flege, J. E. and Bohn, O.-S. (2021). *The Revised Speech Learning Model (SLM-r)*, page 3–83. Cambridge University Press, 1 edition.

Gillette, J., Gleitman, H., Gleitman, L., and Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, page 42.

Goksel, A. and Kerslake, C. (2011). *Turkish. [electronic resource] : an essential grammar.* Routledge essential grammars. Routledge.

Grimaldi, M., Sisinni, B., Gili Fivela, B., Invitto, S., Resta, D., Alku, P., and Brattico, E. (2014). Assimilation of l2 vowels to l1 phonemes governs l2 learning in adulthood: a behavioral and erp study. *Frontiers in Human Neuroscience*, 8.

Kandhadai, P. and Federmeier, K. D. (2010). Automatic and controlled aspects of lexical associative processing in the two cerebral hemispheres. *Psychophysiology.*

Kartushina, N. and Frauenfelder, U. H. (2014). On the effects of l2 perception and of individual differences in l1 production on l2 pronunciation. *Frontiers in Psychology*, 5.

Kutas, M. and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event-related brain potential (erp). *Annual Review of Psychology*, 62(1):621–647.

Kutas, M. and Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science (New York, N.Y.)*, 207(4427):203–205. Citation Key: doi:10.1126/science.7350657tex.eprint: https://www.science.org/doi/pdf/10.1126/science.7350657.

LaCross, A. (2015). Khalkha mongolian speakers' vowel bias: L1 influences on the acquisition of non-adjacent vocalic dependencies. *Language, Cognition and Neuroscience*, 30(9):1033–1047.

Lange, K., Kühn, S., and Filevich, E. (2015). "just another tool for online studies" (jatos): An easy solution for setup and management of web servers supporting online studies. *PLOS ONE*, 10(6):e0130834.

Lei, D., Liu, Y., and van Hell, J. G. (2022). Novel word learning with verbal definitions and images: Tracking consolidation with behavioral and event-related potential measures. *Language Learning*, page lang.12502.

Liu, Y. and Hell, J. G. (2020). Learning novel word meanings: An erp study on lexical consolidation in monolingual, inexperienced foreign language learners. *Language Learning*, 70(S2):45–74.

Lopez-Calderon, J. and Luck, S. J. (2014). Erplab: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8.

Mangardich, H. and Sabbagh, M. A. (2022). Event-related potential studies of cross-situational word learning in four-year-old children. *Journal of Experimental Child Psychology*, 222:105468.

Martin, A. and Peperkamp, S. (2020). Phonetically natural rules benefit from a learning bias: a re-examination of vowel harmony and disharmony. *Phonology*, 37(1):65–90.

Martin, A. and White, J. (2021). Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry*, 52(1):227–239.

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2):314–324.

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7(7):703–704.

Mestres-Misse, A., Rodriguez-Fornells, A., and Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 17(8):1858–1866.

Mulak, K. E., Vlach, H. A., and Escudero, P. (2019). Cross-situational learning of phonologically overlapping words across degrees of ambiguity. *Cognitive Science*, 43(5).

Nobre, A. C. and McCarthy, G. (1994). Language-related erps: Scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience*, 6(3):233–255.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011:1–9.

Osterhout, L., Pitkänen, I., McLaughin, J., and Zeitlin, M. (2019). *Event-related Potentials as Metrics of Foreign Language Learning and Loss*, page 402–416. Oxford University Press.

Perea, M., Hyönä, J., and Marcet, A. (2021). Does vowel harmony affect visual word recognition? evidence from finnish. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Perfetti, C. A., Wlotko, E. W., and Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6):1281–1292.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3):510 – 532.

Rugg, M. D. and Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6):251–257.

Scharinger, M., Idsardi, W. J., and Poe, S. (2011). Neuromagnetic reflections of harmony and constraint violations in turkish. *Laboratory Phonology*, 2(1).

Scharinger, M., Lahiri, A., and Eulitz, C. (2010). Mismatch negativity effects of alternating vowels in morphologically complex word forms. *Journal of Neurolinguistics*, 23(4):383–399.

Skoruppa, K. and Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, 35(2):348–366.

Smith, L. and Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3):1558–1568.

The MathWorks Inc. (2022). Matlab version: 9.13.0 (r2022b).

Tuninetti, A., Mulak, K. E., and Escudero, P. (2020). Cross-situational word learning in two foreign languages: Effects of native language and perceptual difficulty. *Frontiers in Communication*, 5:602471.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., and Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, 68(1–2):306–311.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, 107(2):729–742.

Ylinen, S., Huuskonen, M., Mikkola, K., Saure, E., Sinkkonen, T., and Paavilainen, P. (2016). Predictive coding of phonological rules in auditory cortex: A mismatch negativity study. *Brain and Language*, 162:72–80.

Yu, C. and Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18(5):414–420.

Yurovsky, D. and Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, 145:53–62.

Appendix A

Picture Referents

Real object pictures that were used in the learning phase as referents of pseudowords. The pictures of the least were taken from the Bank Of Standardized Stimuli (BOSS) by Brodeur et al. (2010).

sVbV

kVvV

hVnV

tVzV

pVmV

şVdV

çVgV

fVyV

Note. Vowels (V) in the pseduwords changed according to the vowel harmony conditions

Appendix B

SJT - Auditory words

Table B.1: Novel pseudowords used in Experiment 2

Backness Harmonious		Backness Disharmonious	
Flatness Harmonious	Flatness Disharmonious	Flatness Harmonious	Flatness Disharmonious
pime	çugo	pume	sıbö
fiye	huno	fuye	kıvö
şıda	küvö	tüza	çigo
tıza	sübö	şüda	hino

Table B.2: Real word used in Experiment 2

Semantic category	Real Word
Furniture	Gardırop
Vegetables	Bamya
Clothing	Palto
Trees	Ladin
Part of the Human Body	Bacak
Fish names	Kefal
Flowers	Zambak
Precious Stone	Safir

Appendix C

SJT - Orthographical words

Table C.1: Real word presented orthographical with pseudowords in SJT

Four-Footed Animals	Musical Instrument	Fruits	Vehicles
Köpek	Gitar	Elma	Uçak
Kedi	Flüt	Armut	Dolmuş
İnek	Davul	Kiraz	Tren
Koyun	Keman	Karpuz	Kamyon
Eşek	saz	Erik	Taksi
Keçi	Org	Üzüm	Vapur
Öküz	Zurna	Çilek	Gemii
Tavşan	Kaval	Vişne	Metro
Kitchen Utensil	Baverages	Type of Birds	Carpenter's Tools
Kaşık	Kola	Serçe	Pense
Tabak	Ayran	Kartal	Çekiç
Bıçak	Çay	güvercin	Çivi
Çatal	Bira	Karga	Vida
Bardak	Kahve	Şahin	Keser
Tuzluk	Rakı	Leylek	Testere
Tava	Süt	Bülbül	Matkap
Kepçe	Gazoz	Martı	Makas

Table C.2: Real word presented orthographical with real words in SJT

Furniture	Vegetables	Clothing	Trees
Koltuk	Ispanak	Gömlek	Çam
Masa	Domates	Etek	Kavak
Halı	Patlıcan	Kazak	Meşe
Sandalye	Biber	Ceket	Gürgen
Yatak	Havuç	Tişört	Çınar
Dolap	Marul	Çorap	Söğüt
Sehpa	Kabak	Hırka	Köknar
Perde	Soğan	Yelek	Servi
Part of the Human Body	Fish names	Flowers	Precious Stones
Göz	Hamsi	Gül	Elmas
Kalp	İstavrit	Papatya	Zümrüt
Kulak	Levrek	karanfil	Yakut
Mide	Yunus	Lale	Pırlanta
Ayak	Lüfer	Menekşe	Altın
Burun	Kalkan	Sümbül	İnci
Böbrek	Uskumru	Leylak	Akik
Beyin	Mezgit	Nergis	Gümüş

Appendix D

EEG Pre-Screening Form

Bilkent University

EEG Pre-Screening Form

Participant full name (first name, last name):
Date of birth:
Gender:

General Questions

Right handed? Yes No
Normal hearing? Yes No
Normal (or corrected) vision? Yes No
Native language:

EEG Compatibility

Do you have any psychiatric or neurological disorder (e.g., epilepsy)?

Do you use psychiatric medication?

Do you have any metallic implants around your neck or head?

Are you pregnant or could you possibly be pregnant?

How many hours do you usually sleep?

How many hours did you sleep last night?

Appendix E

Language History Questionnaire

Date ____________ Exp ____________ Participant# ____________

LANGUAGE HISTORY QUESTIONNAIRE

Gender: ____________ **Age(in years):** ____________

Handedness : L/R **Native country:** ____________

This questionnaire is designed to give us a better understanding of your experience learning languages. We ask that you be as accurate and thorough as possible when answering the following questions and thank you for your participation in this study.

If at any time, you need more space to write, please feel free to use an extra sheet of paper. Please put the question number beside your responses.

1. Do you have any known visual or hearing problems (corrected <u>or</u> uncorrected)?

____

____

2. What is your first language (i.e., **language first spoken**)? If more than one, please briefly describe the situations in which each language was used.

____

____

Which language (if any) do you consider your *second* language?

____

3. If you have ever lived in or visited a country where languages other than your native language are spoken, please indicate below the name of the country (countries), the duration of your stay in number of months (if less than 4 weeks, do not include), and which languages you used while you were in the country (please indicate if you were spoken to in a language other than your first language, even if you never actually spoke that language).

Country visited	#Months there	Language(s) used

4. List below, from most fluent to least fluent, all of the languages to which you have been exposed. Also specify the age in years at which you began to learn the language and the context in which you learned it. For example, "Turkish, birth, home". Include all languages to which you have been exposed, although you may never have had formal training in them and may not be able to read, speak or write them.

Please remember to list your native language.

Language	Age in yrs.	Learning Situation

5. What languages were spoken in your home while you were a child and by whom?

____

____

6. Please list the language(s) the following people speak.

Parent(s): ______________________________________

Other caregiver(s): ______________________________

Closest friend: __________________________________

7. How many years have you studied your second language? Please indicate the setting(s) in which you have had experience with the language (i.e., classroom, with friends, foreign country...)

Number of years: ______________________________

Setting(s): ____

____

8. Are you currently enrolled in any language courses? (Any course either instructed in a foreign language or designed to teach a foreign language). If so, please list the course number(s) below, along with the title of the course.

Department	Course #	Title of Course

9. What languages **<u>other</u> than your first language** do you <u>speak</u> fluently?

____

10. What languages **<u>other</u> than your first language** do you <u>read</u> fluently?

____

11. What languages **<u>other</u> than your first language** do you <u>write</u> fluently?

____

12. What languages **<u>other</u> than your first language** do you <u>understand</u> when they're spoken?

____

<u>For the next eight questions,please circle the number of your response</u>

13. Please rate your **first language** <u>reading</u> proficiency on a ten-point scale. (1 = not literate, 10 = very literate)

1 2 3 4 5 6 7 8 9 10

not literate ... very literature

14. Please rate your **second language** reading proficiency on a ten-point scale. (1 = not literate, 10 = very literate)

1 2 3 4 5 6 7 8 9 10

not literate very literature

15. Please rate your **first language** writing proficiency on a ten-point scale. (1 = not literate, 10 = very literate)

1 2 3 4 5 6 7 8 9 10

not literate very literature

16. Please rate your **second language** writing proficiency on a ten-point scale. (1 = not literate, 10 = very literate)

1 2 3 4 5 6 7 8 9 10

not literate very literature

17. Please rate your **first language** conversational fluency on a ten-point scale. (1 = not fluent, 10 = very fluent)

1 2 3 4 5 6 7 8 9 10

not fluent very fluent

18. Please rate your second language conversational fluency on a ten-point scale. (1 = not fluent, 10 = very fluent)

1 2 3 4 5 6 7 8 9 10

not fluent very fluent

19. Please rate your first language speech comprehension ability on a ten-point scale. (1 = unable to understand conversation, 10 = perfectly able to understand conversation)

1 2 3 4 5 6 7 8 9 10

no comprehension perfect comprehension

20. Please rate your second language speech comprehension ability on a ten-point scale. (1 = unable to understand conversation, 10 = perfectly able to understand conversation)

	1	2	3	4	5	6	7	8	9	10
no comprehension										perfect comprehension

21. How would you rate your foreign language learning skills? Please circle your response.

worse than average   average   better than average

22. When learning a foreign language, which of the following do you find the easiest to learn? Please rank the following from **1 to 4**. (1 = easiest, 4 =hardest)

Pronunciation: ______________________________

Grammar: ______________________________

Sayings/Expressions: ______________________________

Vocabulary: ______________________________

23. Have you ever been immersed in your second language culture (e.g., living abroad)?

yes   no

24. Please compare learning a second language in an immersion environment with learning one in a classroom environment. (E.g., Which is easier? In which did you learn more?...). Please comment about the difference between your various learning experiences.

____

____

____

25. Is there anything else about your language background that you would like to comment on? Please feel free to make comments about things which were not covered on this questionnaire.

THANK YOU FOR YOUR PARTICIPATION!